DOCUMENT RESUME

ED 022 529 PS 001 082

By-Garfunkel, Frank; And Others
HEAD START EVALUATION AND RESEARCH CENTER, BOSTON UNIVERSITY. REPORT OF RESEARCH, SEPTEMBER, 1966-AUGUST, 1967.
Boston Univ., Mass.
Spons Agency-Office of Economic Opportunity, Washington, D.C.
Pub Date 67
Note-156p.
EDRS Price MF-$0.75 HC-$6.32
Descriptors-AGGRESSION, BEHAVIOR RATING SCALES, CLINICAL DIAGNOSIS, COMMUNICATION (THOUGHT TRANSFER), CONCEPT TEACHING, EMOTIONALLY DISTURBED CHILDREN, LANGUAGE PROGRAMS, MOTIVATION TECHNIQUES, NONPROFESSIONAL PERSONNEL, PRESCHOOL CHILDREN, *PRESCHOOL PROGRAMS, PRESCHOOL TEACHERS, *RESEARCH PROJECTS, TEACHER BEHAVIOR, TEACHER SEMINARS, TEACHING STYLES
Identifiers-*Head Start

This document is a report of research conducted from September 1966 to August 1967 by the Head Start Evaluation and Research Center of Boston University. Eleven studies and projects are reported, many of them in preliminary or incomplete form because either they are ongoing studies or the data analysis is not finished. The 11 studies contain six topics. These topics and the accession numbers of the individual abstracts are as follows: observation of teaching and curriculum (PA 001 250, PS 001 251, and PS 001 252); emotional disturbance of Head Start and middle class preschool children (PS 001 253); perception and social values (PS 001 254 and PS 001 255); teaching programs and learning (PS 001 256, PS 001 257, and PS 001 258); nonprofessional interviews (PS 001 259); and the teacher seminar (PS 001 260). (WD)

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

BOSTON UNIVERSITY

Boston, Massachusetts

HEAD START EVALUATION AND RESEARCH CENTER

Report of Research
September, 1966-August, 1967

Frank Garfunkel
Director

ED022529

PS 001082

TABLE OF CONTENTS

FOLIO:

A. OBSERVATION OF TEACHING AND CURRICULUM

- A-I Teaching Style: The Development of Teaching Tasks
  Dr. Frank Garfunkel
- A-II Observation of Teachers and Teaching: Strategies and Applications
  Dr. Frank Garfunkel
- A-III Observational Strategies for Obtaining Data on Children and Teachers in Head Start Classes, (OSOD)
  Dr. Frank Garfunkel

B. EMOTIONAL DISTURBANCE OF HEAD START AND MIDDLE-CLASS PRESCHOOL CHILDREN

- B-I Primary and Secondary Prevention: Studying Clinical Process and Disturbance with Preschool Children
  Dr. Frank Garfunkel
  Dr. Caroline Fish
  Dr. Pierre Johannet
  Miss Anne Coolidge

C. PERCEPTION AND SOCIAL VALUES

- C-I Perception of Emotion: Race and Sex Differences of Perceivers and Expressers
  Dr. George Gitter
- C-II The Expression of Aggression in Preschool Children
  Miss Lynn Dorman

D. TEACHING PROGRAMS AND LEARNING

- D-I The Effects of a Teacher Developed, Preschool Language Training Program on First Grade Reading Achievement
  Mrs. Sandra Alexanian
- D-II Training for Number Concept
  Dr. A. H. Blum
- D-III Preferences Among Qualitatively Differing Uncertainties
  Dr. D. Mostofsky

E. OTHER PROJECTS

- E-I The Utilization of Non-Professional Interviewers in the New England and Mississippi Samples by the Boston University Head Start Evaluation and Research Program, 1966-1967
  Mrs. Suzanne Clay
- E-II Teacher Seminar
  Mrs. Sandra Alexanian

INTRODUCTION[1]

The research component of the Boston University Head Start Evaluation and Research Center has involved activity in four major areas:

1. Observation of teaching curriculum
2. Emotional Disturbance of Head Start and Middle-Class preschool children
3. Perception and social values of preschool children
4. Teaching programs and learning

The first two areas have been the principal focus of the Center and will continue to receive considerable attention. The Center Director and his colleagues in the Departments of Special Education, Psychiatry, and Psychology have been actively studying teaching and disturbance for some time and have been given strong University support for the establishment of research programs, extensive observational and recording facilities, and a psycho-educational clinic whose staff and students have provided educational and therapeutic services for disadvantaged preschool children and their families for the past six years. The training program of the Special Education Department includes observational seminars which have provided the raw material from which much of the research reported herein has been generated. The connection between observational strategies and disturbed behavior is the driving force that has led to our strategies and our committments.

There are diverse theoretical and practical reasons why teaching is studied, which leads to a variety of methodological approaches. As we have come to see disturbances in children as often being very much a function of situational (interpersonal) variables, it was inevitable that a large amount of clinical and research effort would be expended on the observer and the observed. The question of whether the former or the latter is affected by observational process or its data products is particularly salient when the focal behavior is painful to children and teachers. Along with a scientific interest in constructing theories that will usefully account for social, emotional and cognitive development, there is a categorical imperative about the potentiality and necessity for change. To the extent that children show extraordinary, but systematically varying behaviors in different situations, it can be inferred that such variability should be possible within any one of these situations. Similarly, when participants of an observational seminar show extreme divergence in their perceptions and interpretations of incidents, there is evidence for the potential of corresponding situational and behavioral change.

Thus, for theoretical and practical reasons, this research in observation and disturbance focuses on polraities. Intellectually and emotionally disabled children and relatively stylized teachers and learning situations are the vehicles for studying systemic instabilities, much as Head Start is a vehicle for more adequately dealing with the dilemmas of poverty. Research is aimed at deviant groups both because of social and psychological needs and theoretical considerations. The study of abnormal individuals and groups tests the limits of human potential for development, as well as for pathology. Disturbed Head Start children are a deviant group within a deviant group. Understanding the dynamics of disturbance will certainly clarify the deliterious effects of poverty and of social impoverishment.

---

[1] "The research reported herein was performed pursuant to a contract with the Office of Economic Opportunity, Executive Office of the President, Washington, D.C., 20506. The opinions expressed herein are those of the author and should not be construed as representing the opinions or policy of any agency of the United States Government."

Studies of perception and social values, and teaching programs and learning have been initiated and are continuing. The structure of the E&R Center has permitted articulation between behavioral and pragmatic research.

The fifth folio includes reports of activities which, although ancillary, were an expression of the tactical bias of this Center. In addition to informal contacts with professionals, non-professionals and parents, a formal consultative arrangement was set up whereby teachers, social workers, community aides and parents were invited to discuss their programs with Center staff and to make recommendations regarding researchable topics and priority variables. Because of the tarnish of subjectivity, blatant facts and issues are often, in effect, denied. This results in assumptions of uniformity accross teachers and programs which leads to specious research and evaluation decisions. The strightforward testing and comparing of diverse groups of Head Start and non-Head Start children would appear to be a classis example of holding one variable constant (the test) and letting everything else vary--a pure anti-experimental design. The sterility of tests for other than predictive paradigms is becoming readily apparent. Tests constructors must worship the idol of reliability and implicity repudiate the iconoclasm of change. This is further confounded by experimenter and other reactive effects.

The longitudinal study mentioned in a previous report is not reported because case finding is incomplete. Several major studies will be undertaken with this sample of 59 school aged disadvantaged children who have been followed for the past five years. During the period that this report covers, 43 of 59 children were located and given the Stanford-Binet and a battery of achievement tests. Current school grades for sample children and their siblings were obtained from their schools. Children and their families have been and are continuing to be followed up disregarding their present location. It has become apparent from keeping in contact with mobile families that drawing developmental inferences from cross-sectional studies is precariousand leads to spurious conclusions about stability and change of disadvantaged children. This will be carefully documented during the coming year.

The Research Conference on Interventions with Preschool Disadvantaged Children, October 24-28, 1966, was discussed in the first quarterly report and the proceedings of that conference were distributed during the Spring. It was followed up by a second conference at Temple University which further provided for dissemination and discussion of current research.

A comprehensive report on the Mississippi phase of the evaluation will be forthcoming when the voluminous anecdotal material is transcribed and assembled so that it can be systematically reviewed. Head Start programs, teachers, children, families, and local conditions contrast dramatically with their counterparts in other regions of the country. They deserve more careful and extensive treatment than we were able to give. It makes little sense to include Mississippi Head Start with other programs with considerable attentions to the peculiar conditions in which it has developed.

The work of the Center was made possible by the unstinting cooperation of Head Start personnel throught New England and Mississippi. We have been greatly impressed by the sincerity and tenacity of teachers, social workers, administrators, and parents in spite of the continuing uncertainty that inevitably surrounds a program that is subject to the erratic cross currents of a society's experiment with itself. We, as well as they, have had to continually be faced with impossible situations that could not be put aside for a more propititous time. Emergencies and the exceptions have become the expected. In Mississippi we were observing teachers and testing children in program whose day-to-day operation was threatened by budgeting limitations which seemed both ludicrous and tragic in a country whose inhabitants were living at a subsistance level. The list of irregularities is far more impossible than that of predictables.

The staff of the Center has given outstanding support to the accomplishment of our evaluation task and to the initiation of research activity. They have been continuously braced and inspired by Boston University faculty and administration.

Barbara Rudolph has capably and steadfastly organized and supervised the typing and printing of this report. Sandra Alexanian, Anne Coolidge, and Suzanne Clay have continuously and insightfully developed a coordinated effort that initially consisted of, almost entirely, loose ends. Among many, we would like particularly to thank the following Head Start personnel for their valuable assistance and trust: Reginald Eaves, Boston; Rheable Edwards, Boston, John Flynn, Cambridge; Matthew Goode, Brockton; Reverand Cornelius Hastie, Boston; William McKain, Bolivar County, Mississippi; Jean Vacco, Boston; Aaron Vence, Bolivar County, Mississippi; and Geraldine Withecomb, Hartford, Connecticut. However, teachers, social workers, and parents were the backbone of our efforts, our fatal weakness for the seductiveness of children, not withstanding.

OBSERVATION OF TEACHING
AND CURRICULUM

HEADSTART EVALUATION AND RESEARCH CENTER

BOSTON UNIVERSITY

TEACHING STYLE:
The Development of Teaching Tasks[1,2]

Frank Garfunkel

ABSTRACT

Tasks were developed and presented to Headstart teachers in order to facilitate descriptions of variation in teaching style. Twenty minute samples of teaching according to task instructions, were filmed so that inter and intra teaching comparisons could be carefully analyzed by diverse observers, thus permitting concommittant study of observer(ing) variation. The use of tasks provides sufficient standardization to permit observers to make accurate predictions regarding subsequent task teaching behaviors. Systematic variations in task requirements will provide a basis for studying more and less invariant characteristics of teaching style which will generate variables that intervene between content and methodology, and individual and group behaviors of children.

---

[1] " The research reported herein was performed pursuant to a contract with the Office of Economic Opportunity, Executive Office of the President, Washington, D.C. 20506. The opinions expressed herein are those of the author and should not be construed as representing the opinions or policy of any agency of the United States Government."

[2] Miss Anne Coolidge perceptively and tenaciously assisted in the development and administration of tasks. Professor Alvin Fiering directed the film making with extraordinary sensitivity. Teachers and observers involved played a critical role throughout the project.

PS 001082

TEACHING STYLE:
The Development of Teaching Tasks[1]

Frank Garfunkel

Boston University

INTRODUCTION:

In order to facilitate comparisons between teachers, tasks were presented to teachers to be carried out with their classes. Teachers were directed to teach with given instructions and materials for approximately twenty minutes. These procedures were used to explore the feasibility of designing, using and recording (filming) tasks in order to provide a vehicle for studying contrasting teacher styles. By having each of six teachers use five separate tasks, it was possible to systematically compare teachers across tasks, and tasks across teachers. As three of these tasks were filmed, there is a permanent record of performance which can be used to study related perceptual styles of performing teachers, observing teachers and other observers. Although, for the purpose of this developmental study, the films are, themselves, the data from which such inferences are to be drawn, it is possible to infer reductions and to make consequent, quantitative within and between teacher comparisons.

The focus of these tasks and films has been to internally validate the use of tasks as a viable technique for comparing irreducable components of style by filming and presenting a series of integrated constellations of behaviors. While the biases of the investigators are implicit in the types of tasks selected, the existence of films provide an objective base for distinguishing stylistic variations in types and degrees of control of materials and situations, and in operational definitions of work and play and their relationship to learning.

Problems of external validation have been only partially and informally dealt with by reviewing teachers' performances in three filmed tasks and two non-filmed tasks in order to ascertain whether there is consistency. Put in other terms, given any one task which is filmed and/or anecdotally recorded, can accurate predictions be made about performance of tasks? With a single exception, it was possible to make

---

[1] "The research reported herein was performed pursuant to a contract with the Office of Economic Opportunity, Executive Office of the President, Washington, D.C. 20506. The opinions expressed herein are those of the author and should not be construed as representing the opinions or policy of any agency of the United State Government."

rather precise forecasts as to how teachers would handle subsequent tasks, specifically with regards to type and degree of control, position of teacher in elaboration of tasks and work-play dichotomizing as it was acted out by the children. The relationship of style to behavioral effects in children outside of the classroom was not attended to either concurrently or longitudinally.

RATIONALE

Although it is rather simple to get teacher variation on any number of dimensions, it is extremely difficult to unravel confounding of teacher, child, curricular, and interactional variables. This is particularly true when dimensionality of teaching behavior is constructively linked with child behavior. Too often the abstraction of teaching behavior is temporarily and conceptually disconstinuous with child behavior and it is not surprising that it has been almost impossible to document the relationship of one to the other. Teachers are measured independently of children and children are measured independently of teachers. This leaves an unknown and unseen terrain--a kind of black box--in which input and output can be documented but for which interaction is ignored. Generally, interactions are attended to separately (Flanders, 1965) and are not necessarily a part of the design that aims at qualifying output by carefully considering the nature and extent of interaction. Furthermore, it is not all clear as to whether given interactional patterns are cause or effect or whether they have been generated because of traditional or teleological determinents.

The theoretical-methodological problem revolves around the choice of an operational structure of variance and invariance. This pertains not only to sample selection--grade level, subject matter, teacher characteristics--but also to measurement strategies--achievement tests, performance in another situation, observational protocols, tape recordings, films. What is to be held constant? What will vary? How will the behavior be recorded and then measured? How will the connection between teaching and learning be established and cross-validated so that we can logically deduce one from the other?

Although holding grade level and subject matter constant would appear to be a useful quasi-experimental device (Bellack et al, 1966), it is necessary to be aware of an implicit trap in this procedure. It may very well be that grade level and subject matter invariance are trivial with regards to teaching-learning variation. Perhaps there are teachers in different grades and subjects that represent a far more critical type of invariance--one that is connected with motivation, transference, values, and creativity. If this is so, then our inferences from grade and subject controlled studies will be only trivially related to either situation or individual outcome. This is by way of saying that holding grade and subject constant does not guarantee that resulting variation of either situations (classroom behaviors) or children (achievement tests) will be meaningfully related to any given dependent variable. However, in spite of this caution, the disregarding of grade and subject would necessarily encumber the comparability of observations. What is needed is control for more obvious independent factors and also control on teaching behavior variation that is more directly related to child behavior over time.

It is not at all clear as to what an optimal segment of behavior might be for any given study. Some studies focus on highly specific signs or categories of motor and verbal activity (Medley and Mitzel, 1963). Participant observation studies consider institutional behaviors over time without specific reference to fragments except as they relate to the developing hypotheses of observers (Becker, 1952, 1953). One study of classroom language behavior used transcriptions of social studies classes and content analyzed thematic material (Bellack, et al, 1966). Critical incident techniques define episodes as they take place in classrooms and analyze them with respect to a field of forces operating in the class (Flanagan et al, 1958). Operant procedures have been applied to teacher and child behaviors in order to study highly specific child-teacher contingencies (Haring and Lovitt, 1967). Each of these strategies is, at the same time, trying to more adequately understand teacher and child in classroom situations and effectively deal with that behavior. The theoretical system from which each is derived is not as important as the logical construct of which it is a part. Method and definitions of behavioral units implicity reflect educational values towards intervention and change.

The several measurement strategies mentioned above vary as to whether or not direct manipulation is involved, and there is a direct connection, over time, between teacher (or class) behavior and individual child behavior. In all of these examples, however, there is either an implied or explicit dependent variable. Classroom behaviors have been studied in order to determine variable effects naturalistically or as a result of specific manipulations. As the desirability of any posited effects is, necessarily, moot it remains to judge strategies either purely in terms of pedagogy or in terms of diverse effects on children over time, including transference of obtained behaviors to other times and situations. Too often, pedagogic variation is buried because of methodological problems in obtaining consistant data. It is as difficult to identify competence in teachers as it is in personalities. Specific performance criteria do not hold up either concurrently (consensus comparisons with other teachers by skilled raters) or longitudinally (relating teacher characteristics to differential achievement performance of children). The failure to identify competence can be partly attributed to several sampling and methodological problems. The homogenizing effect of procedures for selecting and retaining teachers, children and curricula might contribute to the reported error of incorrectly accepting the null hypothesis. For example, if unusually competent and incompetent teachers are eliminated from a sample,the variance will be restricted and differences needed to reject the null hypothesis will be excessive. Furthermore, the pairing of teachers with children of different social classes and abilitites is highly selective, as is the placement of children within schools. Finally, commonly used tests have been developed on the basis of principles and goals that are, in general, at variance with those of intervention. Items that are sensitive to differential treatment appear to be unreliable and are, therefore, eliminated. Probably the strongest single factor that effects item selection for achievement tests is very much developmentally oriented--which results in items being highly correlated with chronological age. If, as it would appear, most measurements used are heavily weighted in this direction, it is unlikely that they will reflect differences due specifically to teachers or, in general, to interventions. Thus,in eliminating items which would tend to reflect day to day subject variation, likely indicators of other sorts of variation are also purged.

The teacher tasks to be described herein have been developed to speak directly to questions of competency and change, both in children and teachers. In order to deal with the relation of competency to change over time with respect to transfer, it is first necessary to determine the nature of situational presses on children and the effect of these presses, if any, over time but within confines of evolving situations. The first question is conditional--if a given teacher (class) has an effect on a particular child, what is it most likely to be? The distinction between the teacher's effect on the class and the differential effects of the class on individual children must be made. But it is doubtful that the latter will be apparent and measurable unless the former is carefully described in terms that cover a broad range of variation. This should eventually provide a basis for dealing with the ultimate question which must be asked about any sequence of behavior that takes place in a class--how appropriate is it for individual children that are exposed? It is not enough to describe the various methods and materials used, nor even the ways in which they are applied. Eventually, attention must be directed to the quality and substance of interventions that children are involved in during the course of the school year.

The use of twenty minute tasks units provides samplings of behaviors that cover reasonably concise cross sections. These include presentation and distribution of materials, implicit or explicit instructions, development and facilitation, transitions, and denouments. The common task across teachers minimizes the difficulty of focusing upon teaching and reaction variation. This is the same rationale for the development of any standardized procedure.

An important variation in individual testing procedures is the extent to which they call for more or less convergent or divergent responses. If teaching tasks were developed to be related to a training program that called for predetermined criterion responses, it would be possible to design "objective" task presentations and scoring procedures. Such criterion responses have been developed at the University of Kansas and have been reported in thus far unpublished manuscripts.

The alternative is to vary tasks and evaluation procedures along the "projective" (divergent) end of the "projective-objective" continuum. Teachers are given stimuli in the form of materials and rather open-ended instructions, much as an individual subject is given a series of Thematic Aperception tests cards. Instead of a strictly verbal response, the teacher gives a complex behavioral response over a designated time and space interval. Themes can be inferred from films or anecdotal records of task responses. It is also possible to use direct behavioral recording or rating scales in order to compare teacher over tasks and teachers over a single task. Contrasts can be facilitated by varying amount of structure in directions and content of task, selection of teachers with greater and lesser stylized approaches to teaching, time between teachers receiving instructions and performing task (latency), age and characteristics of children, history of class, and available physical facilities. The accomplishment of tasks with teachers that have relatively similar groups of children, physical facilities and group history, with systematic variation over content and latency of tasks would be the ideal way to bring stylistic variations of teaching into relief.

TEACHING STYLE

Dimensions of style are measurable variations in approaches used in teaching, rather than what is taught or, strictly speaking, the methodology used. Style, if properly conceptualized and operationalized, will vary over teachers, but will be invariant over groups, content and methodologies. It is, of course, possible that requirements of these aspects of teaching situations can be so stated as to imply style. However, the usefulness of the proposed model requires that overlap be minimized. Variables of style can then be conceived as intervening between inputs (class and teacher history, content, methodology) and outputs (effects on children, teachers and situations.)

Style must be inferred from the daily confrontation that takes place in classrooms. There should be little question about the dynamics of reactiveness that leads to any given confrontation. Teachers' personalities will effect choice of methodology and content which will, in turn, be effected by teachers' reactions to groups of children and supervision. It is postulated that while a given teacher's methodology will vary over time and situations, style will remain relatively constant even if an aspect of style is predictable erraticism. This is analogous to the construct of "cognitive style" as it has been recently articulated in developmental literature. Teaching style differs in that it must be inferred from classroom (interpersonal) situations. However, it is not group interaction analysis nor a study of pedagogical techniques, both of which are subject to variation having to do with immediate environmental demands.

Classrooms develop personalities or temperaments with more or less superficial components. Style focuses on components that are a function of teacher variation, which are relatively stable. Definition and description of style can only come about with systematic variation of non-stylistic factors. The residue of between teacher variance will provide the ultimate source for hypothesized domains of style. These must be further modified by the response variation which can be broadly conceptualized as participation and interaction. These will, as has been previously stated, further effect style, which will be a continuing series of response sets on the part of the teacher. The extent to which stylistic variation can be empirically partialed into relatively independent dimensions is moot. It is just as likely that an ipsative approach would be more appropriate. This would lead to a factorial study of teachers rather than scales. Methodologically, this would call for multi-task studies of teachers so that their ability to deal with a variety of situations would lead to detailed assessment of intra-teacher variation, which is required for this approach. Tasks would be designed in such a way as to systematically underline expressive reactions to content and behavior so that characteristic and reliable mappings could be generated. It is essential that this strategy should not depend upon linearity and additivity unless obtained data is consistent with these assumptions.

Style is inferred from the behaviors of teachers and children in classroom activities. There is no presumption that any particular teacher-child interactions occur, only that classroom activities reflect style by constellations of individual and interactive behaviors. In order to distinguish these constellations a number of behavioral scales have been conceptualized and variously operationalized either directly by behavioral recording, indirectly by utilizing rater judgements or complexly by inferring ratings from sequences of more and less discrete responses of teachers and children. These scales are listed in tabular form as follows:

Scale Name	Basis of Recording	Polarities	
1. Control	Direct	Teacher	-Child
2. Approach	Complex	Punitive	-Supportive
3. Value	Complex	Work	-Play
4. Warmth	Judgemental	Rejecting	-Accepting
5. Humor	Judgemental	Bland	-Humorous
6. Flexibility	Complex	Rigid	-Adaptive
7. Direction	Direct	Aimless	-Purposeful
8. Differentiation	Direct	Undifferentiated	-Individualized

These scales are a first approximation of relevant components of classroom atmosphere. The goal of the measurement procedure is to describe selected aspects of teaching-learning situations as they evolve in more or less structured tasks. Relevance of particular scales to any given tasks will be a function of the demands of both tasks and teachers. Therefore, additional scales will be developed to obtain data on different tasks. While the importance of any single scale will depend on task requirements, it will also be a function of the behavior being studied--stylistic variation. Admittedly, the procedure of allowing criteria to be a function of behavior is complex and somewhat tautological, but to hold criteria constant would lead to the collection of reliable, but irrelevant data. Teaching is certainly not as simple and uniform as straight forward normative measurement procedures would imply. Measuring style has to come to grips with dilemmas of nominal scaling before ordinal comparisons can be meaningful.

Although no definitive position can be taken with regards to the most effective level of abstraction to be used to most validly differentiate teachers, the development and use of these scales and tasks has involved a strategy that calls for maximum reliance on the experience, training and intelligence of observers and the development and use of contrasting tasks used in objectively different ways, rather than on the careful and restricted definition of items which requires only that observers be trained in a particular methodology. "Objective" tests are always restricted by item format and sampling--a restriction which often leads to objectivity only with respect to scoring. Similarly, category and sign systems used for direct behavioral recording are, in general, objective or reliable only in so far as data collection is concerned. Whether these methods permit objective (valid) sampling of content is open to serious question. The limitations of reliable procedures can severely determine the size and type of behavior that is recorded. Scoring reliability requires mechanical procedures so that least-common-denominator observers and definitions can be used, whether or not they are appropriate.

Measurement calls for a series of controls so that confounding of different varying elements can be minimzed. These controls can pertain to observers, situations, or procedures. The above scales call for situational control (common tasks) and observer control (measurement of observers) which, if successful will allow for procedural simplicity in order to minimize the need for control.

Scales were selected deductively by experienced and trained observers as a result of a series of formal and informal exchanges over nominal comparisons between teachers. After extensive observation of forty Headstart classes, observers were required to rank teachers according to whether they liked or disliked their style. The initial nominal distinctions conceived of each observer as representing a different point of view. Extracting the dimensions of observer variation gave a first approximation of nominal style variation. (As the number of observers was small (six) the data was treated informally.) Selected scales appear to best differentiate observer-teacher (object-subject) co-variation. This is to say disregarding whether observers liked or disliked styles of particular teachers, the scales best differentiates observers' judgements in terms of values inferred from their selection of effective (good) teachers. While most measuring instruments are static in that their published format remains unchanged through repeated usage, these scales are meant to be part of a change process. Accomodation and revision will be expected and included as an inegral part of the methodology. In light of this, specific operational definitions would be fatuous. A more realistic operational approach calls for exploring variation as a reciprocal function of definition.

Control depends on teaching, but it also is a function of the values, sensitivity and perspective of observers. Definition which relies on either teaching or observation alone will produce artificial boundries that isolate trivia. A preferable strategy is to communicate the meaning of scales by raising questions about behavior which can be addressed both to teaching and to observers. For control, who controls (or should control) the selection, initiation, continuation, and termination of activities and interactions? and is mastery only a question of skill, cognition and perception or does it not also include control rather than dependency?

Every scale modifies every other scale. When control involves teacher-child interaction, does approach consist of reinforcement or sanctions and are they punitive or supportive? Is it carried out with humor or blandness? These are not questions that lend themselves to the specific and arbitrary behavioral definition that is necessary for wide standardized application. An understanding of humor in teaching will be obtained intensively by teams of observers who struggle with their differing interpretations with suitable procedures and recording equipment. Standardization will be validly obtained only when an explicit reflection of value conflicts is built into structured variations of procedures and definitions. There must be agreements to disagree so that legitimate points of view can contribute to methodological differentiation. This will lead to variations in not only definition of size and type of behavioral units, but to depth of focus as well. Direction can focus on consitant and sustained use of materials but it can also aim at social-emotional interactions of children and/or adults.

Scales are theoretically independent of each other. Therefore, refinement and elaboration will depend on locating teachers and observers who have relatively unique profiles. Inter-scale correlations obtained on these and similar scales range between .50 to .70 (2/3 of correlations) and agreement on ratings of each teacher and scale are all between .50 and .60. Scale ratings uniformally correlate with total scale score between .75 and .90. Overall agreement on mean ratings on all scales ranges between .70 and .80. Thus, there is a fair amount of scale independence but it would appear to be partly a function of observer-rater variation, or, perhaps more accurately, of operational ambiguity which, as has been suggested, is vital to this methodology.

The development and use of these scales is explicitly tied to the task strategy which includes the recording (film) of samples in order to allow for concommittant studies of rater variation. Their use in unrestricted situations where materials and methodologies are fortuitous will confound observer and teacher variables, and resulting data will necessarily be suspect. Furthermore, the use of tasks (and the accompanying scales) assumes considerable knowledge about developmental levels of individual children and, particularly,the existence of intellectual and emotional disturbance of children in the classroom.

PROCEDURES

Thirty tasks were developed by six observers, each of whom had considerable prior experience as teachers of preschool children and as observers of Headstart classes. Each task consisted of listing of necessary materials, procedures, instructions to be given teachers, rationale and method outcomes. Particular emphasis was given to eliciting variation in teaching behavior along the scales of control, approach, and value. Many of the tasks were reviewed with a group of Headstart teachers who had volunteered to take part in a pilot project which would include extensive observation of classes, anecdotal recording of snack time behaviors, trying out selected tasks and filming three tasks for each teacher. The group of six teachers agreed to try two of the film tasks with a two to three day delay between getting instructions and doing tasks, and one of the tasks with no delay--the instructions and materials were presented and the task was done immediately thereafter.

This resulted in eighteen 20-minute filmed tasks, three each for six teachers and their classes. Because of the considerable expense involved, initially only two of these films were processed so that they could be widely shown and plans could be thoughtfully made about processing any or all of the other sixteen films. All films were reviewed in their unprocessed state (sound and picture on separate tapes). As a result, twelve films (six teachers, two tasks) are being processed.

Anecdotal reports were written for all filmed tasks by an observer in the classroom at the time of filming, and for two additional tasks, including a snack time for each of the six classes.

Data comparing teachers, tasks and teacher-task interactions will be obtained when films are ready. However, the purpose of this pilot project was to develop, use and film tasks in order to demonstrate their effectiveness in showing stylistic variation. The data will be used to communicate dimensions of variation that are shown by the films, not to directly describe variation. In a very real sense, the films are the data. Validity will be ascertained by comparing individual teachers on several tasks and films to anecdotal reports.

At each phase of the pilot project, observers were required to make general (scales) and specific (behavioral) predictions about teaching in succeeding tasks. For five of the six teachers the predictions accurately forcasted ensuing tasks, both generally and specifically. The procedure used was relatively crude because of the great stress on task development and the limited number of teachers. However, the success of the procedure suggests that the number of tasks needed to characterize the style of a given teacher will be a function of the accuracy with which predictions can be made of each succeeding task. Predictive efficiency could also be a criteria for the desireability of filming particular teachers and tasks.

TASKS

Proposed tasks included games, construction with different kinds of materials, science, language, color discrimination, music, fantasy activities, and food preparation. They range from completely open ended activities ("do anything you want with these materials") to highly structured sequences. Some of the tasks implicitly or explicitly involve conflict (cops and robbers, not enough materials for all children) and others were directly concerned with language, discussion, story books, verbal games. Some were directed at small groups and others at the entire class, including other adults.

In selecting tasks an attempt was made to use materials and activities that, although familiar, were not commonly used on a daily basis. The exception to the latter condition was the use of snacks as a task. In order to get a first approximation of stylistic differences, the first (and non-filmed) "task" involved observing each class during their snack period. First predictions followed from this.

Four other tasks selected, three of which were filmed, included the following:

1. Masks
2. Balloons
3. Games
4. Homes and families

Complete descriptions of these tasks are included in the appendix.

Tasks were selected to include a variety of curricular dimensions. Masks would obviously evoke fantasy and also a distribution problem as only six masks were given to each class. Balloons provided all children with uninflated balloons and was a relatively play oriented task with the possibility of dealing with scientific applications. Games included instructions for teaching children to play baseball and dealings with competition. Homes and Families involved a discussion with the children with the opportunity of elaboration and interaction. Snacks gave the opportunity to compare classes on an established routine.

All tasks offered opportunities for diverse styles of controlling materials and activities, approaches for supporting or punishing behavior and expression of values with regard to teaching and learning. Although it was assumed that content was theoretically trivial to the purpose of the study it was realized that some tasks would be more provocative, which was desireable because of the goal of getting a characteristic expression of style in a relatively short amount of time.

FURTHER DISCUSSION

The limited use of the films (because of necessary delays in processing) have shown them to be critical to the careful study of stylistic differences between teachers. This will be set forth in considerably greater detail when processed films are intensively compared by observers representing diverse schools of thought with regards to teaching preschool children. At the same time that teaching styles are being compared it will be necessary to study observer variability. It should be possible to weight ratings on style components according to observer characteristics as inferred from observer reactions to contrasting filmed tasks. In this way, a limited number of films of diverse teachers performing a variety of tasks will facilitate comparisons of a larger number of teachers performing non-filmed tasks.

Of the several components of teaching situations it would appear that teaching style is, at the same time, the most difficult to study and the most critical. Although there are practical limitations, content exposure (curriculum) and responsiveness (participation) can be more or less directly measured by time sampling procedures. The amount of time individual children are talking, painting, dancing, and answering questions can be accurately, even if tediously recorded. Similarly, sustained activities can be classified and quantified. But the way or style in which this occurs is critical to consequent values and dispositions of children. How much children have learned from a school experience is not enough. It is critical to find out and describe how they have learned and how they will approach new learning situations. Even though recall and recognition might be useful indicies of transfer, they are, at best, indirect and often misleading. The combination of convergent accumulations of facts with exposures to determinable teaching styles should provide a more powerful estimation of how children,with equally determinable cognitive styles, will be able to deal with future teaching situations, again with more or less determinable styles.

REFERENCES

Becker, H.S., The career of the Chicago school teacher. American Journal of Sociology, 57, 470-477, 1952.

Becker, H.S., The teacher in the authority system of the public school. Journal of Educational Sociology, 27, 128-141, 1953.

Bellack, A.A., Kliebard, H.M., Hyman, R.T. and Smith, F.L. Jr., The Language of the Classroom. New York: Teachers College Press, 1966.

Bruyn, S.T., The Human Perspective in Sociology--the Methodology of Participant Observation. Englewood Cliffs, New Jersey: Prentis-Hall, 1966.

Flanagan, J.C., Pumeroy, S.S. and Tuska, S.A. A new tool for measuring children's behavior. Elementary School Journal, 59, 163-166, 1958.

Flanders, N.A. Teacher Influence, Pupil Attitudes and Achievement. U.S. Department of Health, Education and Welfare, Office of Education, Cooperative Research Monograph Number 12, Washington, D.C.: Government Printing Office, 1965, 126 pp.

Haring, N.G. and Lovitt, T.C. Operant methodology and educational technology in special education. In Haring, N.G. and Schiefelbusch, R.L. (Eds.) Methods in Special Education. New York: McGraw-Hill, 1967.

Medley, D.M. and Mitzel, H.E. Measuring classroom behavior by systematic observation. Handbook of Research on Teaching. Edited by N.L. Gage. Chicago: Rand McNally, 1963. Chapter 6, pp. 247-328.

Appendix

DESCRIPTION OF TASKS

Masks

1. Materials:

   3 black masks
   3 white masks (standard Halloween type)
   6 elastics (unattached)

2. Instructions:

   Give the 3 black masks to a boy in your class. Give the 3 white ones to a girl. Do not specify how they are to be used, but only give the children some indication of how they might be worn by saying "these things are called masks and one way you can wear then is in front of your face and over your eyes" (you may demonstrate if you wish). The only children you must say this to are the two to whom you give the masks. Let the children develop any play or game that you feel is appropriate. If possible, let the children have these for at least one half hour.

   Pick any boy and girl that you wish (if you do it beforehand make a second choice in case of absence) and we would like you to check a list of adjectives for the ones that describe these two children the best. The list will have such words as "timid, talkative, active, sullen, etc" and it may be checked at the end of the morning after the task has been given.

3. Questions:

   How does teacher deal with fantasy and agression?
   Who controls materials and activities and how is this control handled?
   How does teacher and class deal with a situation where there are not enough materials to go around?
   How much structure is presented to children?

Balloons

1. Materials:

   Small balloons of various colors-one for each child in the class and 4 extras

2. Instructions:

   Leave approximately 20 minutes of your schedule open for this activity. This time allowance is just to give you some idea of how much time this activity might take so that you can get it into your schedule. Please do not feel bound by this-take more or less time as you feel is needed.

   Bring the children together in a group on the floor in a large open space. (If necessary, please push the furniture and equipment to the edges of the room to allow for a large and open space.) Tell the children, "I have one balloon for each of you to play with today. We have lots of space around us here so that you can play with your balloon in any way you would like. You can move with it, hit it up in the air, or do anything else you would like with it in the next 20 minutes." Add any other directions, suggestions, or comments to the children that you feel would be helpful or necessary but be sure to include the above statements. Give each child a deflated balloon. You will be provided with one balloon for each child in your class plus a few extras in case there are any balloons with defects or in case any become broken in the process of trying to blow them up.

We would like you to remain available to the children during these 20 minutes but we do not have anything particular planned for you to do during this time so feel free to either participate, direct, or observe as you would like.

You may dispose of the balloons as you would like after the 20 minutes. For example, the children can take them home or you can keep them at school, etc. If any individual children spontaneously ask during the 20 minutes if they can take the balloons home, answer yes or no as you have decided but please don't announce this fact to the group until after the 20 minutes is over.

Questions:

A. Teacher's ability to anticipate and handle frustration. 1) Does she expect the children to be able to blow up the balloons themselves? Does she anticipate that some children won't be able to do this? How does she handle the frustration of the children who can't blow their balloons up? 2) Does the teacher anticipate frustration from balloons popping and children not being able to have another? How does she prepare the children for this and how does she deal with it afterwards? 3) How does she handle the choosing of colors?

B. If many children ask for help blowing up their balloons, how does she handle them? e.g. Does she encourage the children to try to do it themselves? Does she blow it up for them? Does she ask the aide to help, too? Does she suggest the children ask each other for help? Does she announce to the group that she and the aide are available for help or does she wait for the children to seek out her help?

C. Teacher's ability to handle aggression. Do children try to pop each other's balloons? Do they run into each other? If so, how does teacher handle?

D. What is the teacher's reaction to children whose balloons have popped? Sympathy? "That's life" attitude? "I told you so" attitude?

Homes and Families

1. Materials:

   None

2. Instructions:

   Choose a group of children with whom you will spend a period of time on two successive days discussing their homes and families. We would like to observe during the period inthe morning when you plan to do this. (If the teacher asks about their homes and families just say that anything she thinks would be interesting or good for the children that she is working with.) Write and illustrate discussions with children.

3. Questions:

What aspects of environment or family does teacher focus on or does she let children determine what happens? Does she make any attempt to talk about feelings?
How are children involved in writing and illustrating discussion?
How does teacher react to reports and stories of children?

Games

1. Materials:

Large rubber ball

2. Instructions:

Have children play dodge ball where children are divided into two groups, half inside a circle of the other children. Children in circle eliminate children inside of circle by throwing ball at them.

3. Questions:

How does teacher modify game for children?
How are explanations made?
How are two groups chosen?
How is competition dealt with?

HEADSTART EVALUATION AND RESEARCH CENTER

Observation of Teachers and Teaching: Strategies and Applications[1,2]

Frank Garfunkel

Boston University

ABSTRACT

The rationale for participant observation call for a greater reliance on experience and training of observers and on systematic procedures for sample selection and inter-class comparisons, than on the development of a system for directly and reliably recording categories or signs of behavioral fragments. Variations in teaching and in observation must be analyzed as interdependent sources which both contribute meaningful descriptions of differences between classes. Recording samples of observed behaviors is essential for training and analysis.

Applications using teams of observers in Head Start and inner city and suburban elementary school classes are described and discussed with reference to methodology and data reduction. Films were made of a stratified sample of classes in order to anchor observational reports and ratings and for the purpose of providing primary data on stylistic variation across school location and grade level.

---

[1] "The research reported herein was performed pursuant to a contract with the Office of Economic Opportunity, Executive Office of the President, Washington, D.C., 20506. The opinions expressed herein are those of the author and should not be construed as representing the opinions or policy of any agency of the United States Government."

[2] Films for this project were made under the resourceful direction of Professor Alvin Fiering. Mr. Charles Kokaska performed extraordinarily in developing positive relationships with teachers and in supervising field observers. Miss Janet Hudson has indexed films and organized data with consummate skill.

OBSERVATION OF TEACHERS AND TEACHING: STRATEGIES AND APPLICATIONS[1]

Frank Garfunkel

Boston University

INTRODUCTION

All too often educational studies employ a single recording technique to abstract teacher behavior into data. The monolith is this singular strategy rather than the claims and procedures of any one school of observational thought. Such a criticism is not confined to educational research, but to any studies that focus on complex human behaviors for which there is no optimal methodology that is accepted by professional consensus as being the epitome of validity. Although a particular methodological approach - participant observation (Bruyn, 1966) - will be described, the discussion of perspective is crucial to its elaboration. The vehicle of inference for participant observation is "observer" with experience, training, and theory rather than rating scale, checklist or behavioral protocal. In order to comprehend the validity of any of these vehicles it is necessary to explore their potential diverse contributions and to carefully describe defects in instrumentation, methodology and substance.

Participant observation is not cast as the only or preferred approach, but rather as a necessary component of research activity that aims at inferring useful data from teacher behaviors. The fact that such a strategy does not result in easily reportable and grossly comparable data should not be a deterrrent to its use if there is reason to believe that the behavior being studied is so diverse and complex that descriptive problems are inherent because of this diversity and complexity. Social sciences (and other sciences, as well) always run the rish of reporting that which is easy to describe rather than that which is important to the phenomena being studied.

RATIONALE

Strategies for obtaining data on teacher variation cover a wide range of procedures. Quantification is variously based on rating scales, behavioral categories, checklists, interaction analyses and projective inferences. Reliability is more a question of definition of behavioral units than of their relevance to teacher effectiveness. The substance of the behavior that is designated by the observational model is a reflection of either the instrument maker's or the observer's bias. Whichever

---

[1] "The research reported herein was performed pursuant to a contract with the Office of Economic Opportunity, Executive Office of the President, Washington, D.C., 20506. The opinions expressed herein are those of the author and should not be construed as representing the opinions or policy of any agency of the United States Government."

is the case, there is always a presumption about educational goals and effective implementation. This is just as true of rating scales as of direct measurement which must make a prior decision about what is to be observed. It is not clear that any extant system is based on a theory which would systematically direct us to study particular behavioral categories.

When explicit attempts are made to empirically judge effectiveness by observing changes in children during the time they are with a particular teacher, and, furthermore, to select units of teacher behavior because of their relation to change, there are snarls because teacher effects are engulfed by developmental and social class effects and also, and perhaps more significantly, the behaviors that are most directly affected by teachers are not easily defined or measured. Achievement tests give an abstraction of intellectual behavior which may very well be invariant to teacher effects, especially when compared to intelligence and social class variance. This is not to imply that there are no teacher effects, but only that given the instruments and variables conventionally used, for practical purposes, they are not measurable, at least with the samples of teachers and children that have been used in teacher effectiveness research. This is an important "at least" for, as has been pointed out in psychotherapy research, demonstrated effectiveness of a particular therapist or procedure is very much a function of the diagnosis and severity of the patient. It is possible and probably that teacher effectiveness studies must take into careful consideration the age, sex, and educational-intellectual status of students. The teacher variable will probably prove to be more demonstrably effective for disadvantaged, disturbed, retarded and generally disabled children than for normal children because the variability in criteria is, to a large degree, accounted for by independent variables that are constructively and methodologically highly correlated. There is a confounding between the research problem - are teachers differentially effective? - and the measurement problem, that is largely unresolved.

While admitting that the ultimate criteria of teacher effectiveness are changes in children, it does not necessarily follow that the important teacher variable (or variables) should be derived by regressing changes (in children) against a myriad of input variables (teacher behaviors). For this to be the recommended procedure it would have to be established that the criteria are desirable and that they are meaningfully linked to teacher behaviors, neither of which is definitively so. Research on teaching is faced with a forbidding gap between teaching and learning which is partly a function of the autonomy of teachers and partly of the nature and limitations of teaching and measurement technology.

Failure to develop a predictive system for determining effectiveness has been accompanied by (and partly by default led to) the development of authoritative systems whereby one or more professionals describe what makes an effective teacher. Items, scales or categories are abstracted so that they can be used by a more or less skilled observer, to obtain data on the purported effectiveness of a sample of teachers. Behavioral units can be quite global, encompassing such broad areas as permissiveness, warmth, creativity or control, or they can be extremely specific and relatively nonjudgemental, such as recording the number of times or amount of time that particular behaviors and interactions take place. Global assessment depends on trained and experienced observers while specific assessment depends on trained but not necessarily experienced observers (experience referring to teaching and training referring to observer training).

The construct validity of any more or less global or specfic system will depend on not only the substance of categories or items, but on other disiderata as well. In fact, substance might very well be of least significiance in light of situational and procedural varibilities that are often erroneously assumed to be relatively constant. Given the fact that teachers vary, it does not necessarily follow that procedures are directly comparable, operational goals are the same, samples of children in different classes require the same approach, curricular and time of day variations are insignificant or cultural forces or particular schools are not predisposing. When the burden is on the instrument (rather than the observer) it is difficult or impossible to correct for confounding that is implicit in each of these sources of variation. Given instruments will only be effective to the extent that these intervening variables are not only controlled for (presumably by randomization or manipulation) but are measured and, it follows, whose distributions are adquately represented in the given sample of classes. This suggests that either studies of teaching should concentrate on intensive surveys of relatively homogenious clusters of classes that differ on few but potent dimensions, or that large scale studies include manipulation of curricular, sampling of children, in-service training and supervision. This is to say that there is too much noise in the system for any single instrument to validly assess teacher effectiveness. This is just as true if the instrument is based on a construct as it is if it has been empirically derived.

Another rather imposing source of variation is the observer both-the procedures by which he is trained and those that he uses in the course of his observations. It is not only that different people see different things, but that the conditions of training, visiting classes, feedback, and articulation cannot be assumed to be constant. The use of a single instrument will not insure comparable data unless either the observational process in continuously standardized, the instrument has built in features which suppress observer and observing contamination, additional data is collected to provide for necessary nominal distinctions, or the variability in phenomena being observed dominated observer variability in a direction consonant with the purpose of the data gathering process.

It follows that no single strategy is inherently superior to another one but that there are situational, temporal, economic, and personnel considerations which will suggest that one approach will be more valid than another. The reduction of teaching behavior is desirable because inference is based on more clearly understood judgements. However, reduction can lead to spurious and often misleading data, if it is not accompanied by compatible reduction of other relevant behaviors of teachers, children, and schools. Furthermore, the sin quo non of reduction is that the transformation be reversible. If reduction leads to a collection of irreversible bits that cannot be associated with the child's and teacher's other (and more global) behaviors, then studies of teaching will leave the domain of education and enter some other (possibly meaningful) domain. There are obviously impelling reasons why teaching should be validly assessed, not the least of which is upgrading teacher education, gaining insights into learning of both teachers and children and studying social interactions. If reductionism leads away from these by so abstracting and fragmenting behavior then it is likely that it will contribute much more to behavioral analysis than to change.

The greater the reduction to highly reliable bits of teacher behavior, the more likely it is that accurate predictions will be made of correspondingly reduced to bits of child behavior. Therefore, if the research goal is to get such correspondance, disregarding its relevance for teaching and learning, then maximal reduction is to be desired. But the reduction process, in general, ignores relevance and only accidentally provides indices for units of behavior that are clinically meaningful. Human behavior has not been structured (theoretically) as an accumulation of behavioral bits that go together in an orderly and linear model. It is not at all clear that these bits have any useful meaning by themselves. It is a pragmatic question that can be dealt with only in terms of specified applications which become the guage of usefulness. The research decision to concentrate on any given units is germane not only to methodological considerations--how is the unit best measured?--but to the theoretical connection between teacher and learner. This connection can be conceptualized as being mapped by any level of obstraction or generality. The crucial question arises when clinical requirements demand reversibility--that results under any system of inquiry be useful as feedback in order to affect behavior other than that which is under a microscope. There is just as much need for transfer from datum to person as there is from skill to ability. Without this transfer both systems would be sterile.

Transfer is implicit in a well ordered and predictable system where reversibility (from behavior to abstraction to behavior) is generated from an object (intra) and across objects (inter). An individual's within variability over abilities is reflective of sampling variation across individuals and time and vice versa. The Stanford-Binet IQ is reversible (for middle class children) not because we can go directly back to the individual from the IQ, but because we can go the sample and then, in a meaningful way, back to the individual. "Meaningful way" refers to the well ordered system whereby a probablistic statement can be made about the individual's future academic behavior with regards to the group. Without this characteristic test scores or observational data become one way streets that make no useful connections.

Classroom observation is up against the reversibility dilemma no matter how abstract or reliable are the protocols. When data are obtained they may fit into a regression analysis but they cannot be transformed back to the class either directly or indirectly because of the lack of order in the system, either horizontally or vertically. Because of this,films (or Kineoscopic tapes) are needed to provide a mechanical vehicle for reversibility in the absence of a theoretical or empirical vehicle. Admittedly this only provides for the reversibility; it is not established. But at least the possibility exists. At the same time the vehicle for transfer is present--various techniques can be applied to the same sample of classrooms. Variability of multiple dimensions and strategies can be put to the crude, but immediate test of viewer (film) variability. Direct comparisons can be made between direct recordings of behavioral bits, ratings of qualities, and authoratative judgements. And, most significantly, teachers can be confronted simultaneously with data and behavior. For the present, films would appear to be necessary for the development of any form of observational analysis--without films even carefully obtained data will be lost to a specific, non-transferable and irreversible "black box" process.

The fact that the introduction of the photographer or the observer transforms the situation is not without theoretical interest. If non-reactive procedures can be used in educational studies, as was done by Sexton (1961) and as is recommended by Webb, et al (1966), they are to be desired unless the reactive effects are theoretically important in the reconstruction of phenomena being observed. There is reason to believe that the principle characteristic of teaching is that it is not observed and that feedback is not existent and, in fact, impossible. Education is essentially a nonreactive system which is unaffected by contemporary social movements, recent scientific advances and critical reappraisal of current practices. Authorities set up the models and pontificate but teachers and principals run the show in autonomous conclaves. This autonomy is personal rather than professional. Textbook and examination conformity is obviated by variability along indeterminate and self-defeating lines. The model of classroom observer (or photographer) is one that involves more than an invasion into the classroom for the convenience of research. It is a different and more viable model that permits (but does not insure) a continual reappraisal of curriculum and behavior. The study of unobservable teachers is a paradox without resolution. Teaching conceived as art, science, or some combination of the two is untenable unless it can be researched on the one hand, or experienced on the other. Given the present state of research technology, the falling tree in the forest does not make a sound unless there is someone (or something) to hear it.

Orchestras need listeners, recorders and critics less they exist in an incestuous vacuum. The reinforcement of teachers consists of a bundle of meretricious acts and words which contribute more to a religion than a profession and more to a mystical epistomology than to a vital language that has some relation to behavior. Therefore, the criticism that the observer changes the situation is accepted and encouraged. That the necessary research vehicle is just as essential to pedagogy is not a coincidence. The claim can be made (even if it cannot be rigorously supported) that any social scientific techniques should have direct payoff to the individual or groups being observed and manipulated. Using film to study teaching is an example of this claim.

Disregarding the technique used to record behavior, observational studies are usually confronted by comparisons of teaching that depend on values rather than behavior. If comparisons are to be made between teachers who lecture and those who lead discussions in varying subject fields, any system of measurement will break down unless it is either assumed that one approach is inherently better than the other (values) or that the different behaviors are irrelevant to the measurement of effeciveness which is to assume that goals transcent methodology. There are several ways around this dilemma. The curriculum and/or methodology can be stipulated (Belleck, et al 1966) and teaching can be thusly compared. Unless teachers have opportunities for participation in several manipulations there will be teacher-method confounding. Manipulation can be contrived (with or without teacher involvement) or they can be unobstrusive (and thus really not manipulations) by selecting sequences of comparable behaviors that already exist. In either case and disregarding the observational and recording technique, there is some control so that "everything being equal" is not a completely empty phrase.

If manipulations of the first or second kind are impossible to accomplish, adjustments must be made either by restricting the field of study or by using an "instrument" that allows for diverse methods, curricular and samples. Such an "instrument" might be a series of conditional scales which are selected by the observer depending on the curriculum and techniques being used. Comparisons could be made on those scales that were selected a sufficient number of times. The "instrument" could also be a highly trained and experienced team of observers who have necessary skills to compare somewhat dissimilar teaching situations. To assume, as is often done, that the observer who has the task of selecting and judging, will be more subjective than a series of protocols that cannot deal with the complexities of teaching variance, necessarily involves the tautology that such an observer is definitively subjective, and direct behavioral recording and rating scales are definitively objective. This fallacy is an inheritance of the so called "objective" test which is presumed to be objective because of its format, not because of its item selection, mode of inquiry or reactive effects. Admittedly, the scoring process is less subject to the biases of the scorer and the paper and pencil standardization conditions of test administration are relatively constant, but this does not provide sufficient conditions for objectivity. Reliability is an aspect of what might be referred to as internal objectivity[2] but it is not necessarily primary. It is necessary to consider the effect of the instrument on not only the subject but the educational process, the selection of items, the mode of item presentation and the problems inherent in the transformation of behavior to data. The high reliability of "objective" tests is not without a price in external subjectivity. The assumption that reliability is generic to validity has already been challenged with regard to "objective" testing and it can be similarly challenged with regard to "objective" recording of teaching behavior.

The argument is the same. The selection of items and modes of presentation involves gross subjectivity even though recording and scoring processes (which can be one and the same) are highly reliable. This is not to say that essay tests and the use of the observer-as-instrument necessarily insure external objectivity but only that they provide an alternative strategy which can more directly get at higher level processes. Thinking, reasoning, problem solving and creativity may be vague but they come closer to the expressed goals of education than memorizing, recalling, and educated guessing. Similarly, the assessment of humane, creative, elaborative, insightful, and intelligent teaching is more directly to the point than counting the amount and number of times teachers and students ask questions, make statements, make demands, and are silent. This is not to preclude that specifically defined behaviors can be important indicators of generalized functions but only to gain perspective about their limitations and the value of alternative "subjective" strategies to approach a more profound objectivity than is to be had by using "objective" methods exclusively.

The question of reaction is not a trivial methodological issue that can be relegated to vagaries of research. The teacher who is "counted" and the observer who is counting are part of the system and will respond in some way to this procedure as opposed to an alternative one. The reductionism involved in "counting" reduces not only behavior, but the work and status of the observer and, therefore, of observational process. This is not a polemic for eliminating "counting" but rather an argument for questioning any reactive procedure, not because it is reactive, but because of the quality and force of the reaction it might evoke.

GENERAL STATEMENT OF PROCEDURES

We address ourselves specifically to the problem of evaluating and describing the potential effectiveness of teaching in a diverse sample of classrooms and schools (or centers). Amount of observation will depend on sample vaiability and O sophistication. In order to obtain approximations of these parameters the design calls for multiple O's making multiple observations of classes over an extended period of time. O's will have had teaching experience and will participate in seminars prior to and throughout the PO. Training will consist of a variety of experiences aimed at facilitating inter-O communication, becoming familiar with a behavioral model and developing observation sensitivity. Seminars and workshops prior to PO will be used to screen out unsuitable candidates. O's will participate in an observational seminar where they collectivly observe groups of children in classes and discuss at length, teaching and learning as they view it. O's will observe each other teaching children and discuss varieties of approaches and values.

---

[2] This followed Campbell and Stanley's (1963) distinction between internal and external validity.

Films will be utilized in the observational seminar in order to allow for review of discussed behaviors at any time. These films should show diverse teachers doing similar tasks and similar teachers, or a given teacher, functioning in varying ways. It is desirable for O's to view different teachers with the same group of children.

O's will keep careful logs of observed behaviors which will provide detailed accounts of teacher, child and interactional behaviors. Analytical reports will be written, utilizing the log as sources of evidence. Finally, O's will write interpretive summaries of teachers and classes, describing their estimation of effectiveness and indicating teaching characterisitics that are critical for their assessment. Procedures for writing these reports are set forth in greater detail in the appendix to this report.

Scales representing important and adequately variable dimensions of teaching and child behavior will be constructed in such a way as to relate the observed behavior to the behavioral theory. O's will Q-sort classes on each of these scales--rating all classes on one scale at a time thus minimizing associational biases. O's will underline and label logged behavioral recording according to a notation that related scales to specific recorded behaviors. Scaled judgements can then be supported by molar sequences of observed and recorded behaviors.

MODEL

Although participant observation (PO) varies as to the specific procedures used, it is always based on the principle that although the observer (O) will adapt preconceived structural outlines and dimensional scales on the course of his summary, he is the instrument for inferring data, rather than any outlines or scales. There must be enough intensity and duration in the involvement with the phenomena being studied for its unique structure and process to be indentifiable. The amount of contact is a function of the kind and degree of distinctions between individuals and agencies that are required. Once the target system is defined O has the responsibility of determining a traffic pattern for himself which will lead to an understanding of relationships and direction. Hypotheses are constructed by relating a presumed general theory of behavior to the behaviors of the system. PO methodology is independent of the theory or of the working hypotheses--but some articulated theory is necessary.

O is presumed to be experienced and trained although specifications for both depend on task requirements. Training can be presumed from the previous experience of O or it can take place prior to and during PO. Reliability will depend on the perspective and sensitivity of O and multiple O's can be used to provide anchoring if diverse situations are to be observed. O will observe and become involved (interviews, utilization of unobtrusive data, manipulation) to an extent necessary to test hypotheses about prediced outcomes and structural relationships. Guidelines for participation must be drawn up, prior to observation, with the cooperation of individuals involved.

Biases of O must be continuously dealt with but this will depend on whether they are a legitimate source of error. Where O bias will produce variation equal to or greater than phenomenological variation , it is necessary to articulate and hypothesize about bias x behavior interaction in a manner suggested in general terms by Mydral (1953). Where bias is of minial important (as in many cultural anthropological studies) it need be only articulated.

Just as in any data gathering process, inferences are only as strong as the instruments that are used. PO depends on high quality O's who can demonstrate their perspicacity by being able to predict interactions and circumstances and to relate observed behavior to given theoretical models. Proof of quality can either be left to the readers of final reports or it can be currently brought into relief by using multiple O's with parallel systems. The test of effectiveness or precision is clearly not a reliability coefficient or an "F" ratio. Any such statistical test works smoothly once the data is obtained and disregarding the validity of the data. PO emphasizes letting meaning speak for itself in much the same way the Skinnerians proclaim that data should be directly recorded and then speak for itself.

The assumption of PO is that there are O's and methodologies which can be used to obtain data that reveals more about observed processes than about O's. Methodologies can be designed to efficiently utilize O's with given degrees of competing biases and with specified goals with reference to designated behavioral systems. This is to say that design will have to be adapted for known variations in O's, goals and systems.

PO is not clearly defined methodology that is uniformly used in the social sciences. The practice of having an O look closely at a segment of interpersonal (or individual) behavior is simpleminded and elementary. Where more clearly defined procedures are appropriate they should certainly be used. The designations of adequate O's is difficult and perhaps, often impossible. It might appear that PO is a regression to pre-scientific methodology, where uncontrolled judgements are combined with unknown weights. But is is even less scientific to use "powerful" instruments to perform tasks for which they are unsuited. The decision to use PO is made in light of the complexities of teaching, the difficulties of obtaining comparable samples of behavior, the problems of irreversibility, the tenuousness of child behavioral criteria and the obscurity and ineffectiveness and inappropriateness of personality measurement for obtaining adequate measurements of teacher characteristics. This could lead to the abandonment of such research or, as in the case of PO, to the adaption of relatively crude processes which can, albeit subjectively, deal with those obstacles. Developments in audio-visual technology will make it possible to give more substance to the inferences of O's and to provide reasonably direct documentation of classroom processes that can be exposed to more varified procedures.

APPLICATIONS IN HEAD START AND ELEMENTARY SCHOOL CLASSES

Applications of modified participant observation approaches were made on selected Head Start and Elementary School classes in connection with two projects, which were taking place concurrently. The first involved twenty Head Start classes which were being evaluated by the Boston University Head Start Evaluation and Research Center as a part of its participation in the National Evaluation Program. The second was with Project Concern, an experimental study of the effects of suburban education on inner-city children in and around Hartford, Connecticut. Since the data on both of these projects, with regards to the tested and observed performances of individual children, has not been made available, this report is necessarily incomplete. Procedures for observing classes and obtaining data will be described in some detail, and preliminary descriptive statements will be made with regards to dimensionality of scales that were used in each investigation and agreement between raters on a variety of scale ratings. In addition, for the Project Concern application, the division of classes into inner-city defacto segregated and suburban unite with one, two or three bussed negro children in them permits a straight forward comparison over location of classes.

Although the general principles behind participant observation, as developed by Bruyn (1966) were followed in the development and carrying out of procedures, the sustained and intensive contact of observers with classrooms and schools was not followed partly by choice, because of the kinds of variation that were of most interest, and partly by necessity. Future studies will provide for considerably more contact between observers and the institutions they are observing in order to realize the depth which is only being approximated by procedures to be reported herein.

The aims of these studies were twofold: 1) to study the relationship between selected characteristics of teacher style and changes in mental abilities, academic achievement, personal-social development and creativity of children in selected classrooms; 2) to describe, through cross-sectional procedures, teaching situations which Head Start children are exposed to and those to which they will most probably be exposed to if they attend inner-city or suburban elementary classrooms.

PROCEDURES

Both applications called for the recruiting and training of observers who had extensive experience both as teachers and as observers of preschool and elementary school classes. Initial training sessions involved observation of classes and discussion of an all-inclusive categorical model of classroom procedures (Appendix C). This model was not for the purpose of providing a checklist or of focusing observers' attention on particular variables so much as it was for directing their attention to all possible contingencies and teaching situations. The model included listings, under the general heading instruction, of materials, lessons, motivation, evaluation, and achievement. A second section under the general heading of controls included form, quantity, tone, consistency and student pressure. Facilities listed characteristics, and implications for teaching. Student interaction included opportunity characteristics. A last category, teacher-student interaction included humor, address, feelings, reinforcement. This model was meant to be a vehicle which would serve to provoke discussion and generate questions about varieties of teaching experiences. In addition, an exhaustive list of variables associated with teachers, students and curriculum was constructed, through the deliberations of observers, in order to sensitize them to differences between independent, intervening and dependent variables (Appendix F). It is critical to note that the models developed from observational seminars and were, therefore, the produce of the efforts of observers. They were not handed listings of categories and variables which had been developed externally and which would have been, therefore, imposed upon them.

Observers were asked to keep detailed notes on their observations without regard to a particular model, but with specific regard to what they considered to be the most important characteristics of the classrooms they were observing. These notes were to be transformed into process reports which were to be concluded by analytical reports and summary interpretations (Appendix D).

Scales were developed for both studies by observers after carefully and deductively describing contrasting characteristics of teaching situations which observers judged as being relatively unique. (See appendices A&E). The scales, are, therefore, a reflection of differences seen by observers, rather than the basis for making distinctions. This meant that this approach to studying teaching involved a concomitant study of observer variation, and that these two separate focuses were mutually interdependent.

The burden of responsibility was clearly on observers rather than scales and it called for an inferential process which would be only as defensible as the perceptiveness and intelligence of the observers permitted. This process structures a systematic approach to dealing with subjective impressions of observers who are required to defend these impressions in the face of careful scrutiny by other observers and by senior members of the project staff. The process assumes that each observer has enough experience and insight to be able to produce salient reports and interpretations of teaching variation. Resulting inference must attend to both sources of variation--teaching and observing--in order to adequately describe stylistic variation within stylistic categories.

In order to provide a superstructure for teaching and observing variations, films of selected classes were developed. In covering a wide range of activities, these films have and will continue to provide referent behaviors for the reports and ratings of observers. Extensive use of these films has been and is continuing to be made in order to clarify reductions of behavior that were made by observers.

OBSERVATION OF HEAD START CLASSES

Of the twenty sample classes used in the National Evaluation program, nineteen were observed sufficiently by two or more observers to produce reports and ratings on a series of scales which were constructed by observers during the course of their observations.

Eight scales were used in rating nineteen teachers by six observers, with each teacher being rated by two, three, or four separate observers.

The scales were as follows:

1. Attitude towards teaching situation.
2. Teachers differentiation of children and activities.
3. Predominant emphasis of curriculum.
4. Purposefulness of classroom behavior.
5. Control of materials and interactions.
6. Communication-responsiveness.
7. Work-play continuum.
8. Overall rating.

The detailed statements about each of these scales were given to each observer and can be found in Appendix A.

Rater agreement on the ten scales varied between 80% and 90% and on the overall rating the agreement was 92%. Interscale correlations varied between .60 and .90. Variation between classes appear to be sufficient to allow for maximal rater agreement as well as the probable inflation of scale inter-correlation.

Observers were instructed to sort all teachers on each scale, rather than rating each teacher on all scales, in order to minimize halo effects.

Since four of the six observers had training and experience in early childhood education and, consequently held a point of view which valued highly differentiated programs with a considerable amount of freedom for individual children, resulting ratings are necessarily a reflection of this point of view and are, therefore limited in their generality. Observational teams that participate in such a strategy should represent a wide spectrum of points of view with at least two observers representing each major variation. Similarly,it is essential to obtain samples of classes where competence and style are relatively independent so that their respective sources of variance can be partialed out.

PROJECT CONCERN: Comparisons of inner-city and suburban classes.

Project Concern is a large scale interventional project which provides for educational placement and supportive services for 250 inner-city children. The inner-city children are all residents of Hartford, Connecticut, and the experimental intervention consists of placement in surrounding middle class suburban schools. A randomly selected control group of 250 children is being studied concurrently in order to test hypotheses regarding the differential effects of inner-city and suburban schools on children. A summary of the theoretical framework and the experimental design of Project Concern can be found in Appendix B.

The Boston University Head Start Evaluation and Research Center has been involved in observing and filming a random sample of classes that contain experimental and control children. Observations have also been made on a sample of Head Start classes so that educational continuity between Head Start and elementary school could be ascertained. Filming took place within a careful observational survey design so that the validity of the filming process could be evaluated.

From thirty-nine schools involved in Project Concern, thirty-eight classes were selected for the observational and film survey. Ten of these classes were filmed over a five-month period. The extent to which filmed behaviors of particular classes represent those classes, as well as the extent to which the film classes are representative of all classes, is presently under careful consideration. Findings thus far are that independent observers can go from films to reports and from reports to films with equal facility and that ratings of films are in almost complete agreement with observer ratings made of filmed classes at other times during the year.

Both the observational and film survey included kindergarten, first, second, third, and fifth grades in both inner-city and suburban schools Inner-city classes were selected randomly (stratified on grade) from the total pool of control classes. Suburban classes were selected randomly from two communities that had greatest participation in the project and that represented more and less cooperative coomunities with regards to Project Concern.

The observational team consisted of five observers with widely different backgrounds and points of view. They were trained, respectively, in preschool education, elementary education, elementary and special education, secondary and special education, and elementary education and counseling. Each observer was randomly assigned a sample of classes in both inner city and suburban schools. They were required to make at least two extended observations, separated in time by at least one week, and, preferably, three or four separate observations. In addition, each observer was required to observe classes of two other observers at least once and, preferably, twice each.

Observers wrote process and interpretive reports and rated each class on ten scales that had been derived by the observational team from preliminary observations of the total sample of classes. A sorting technique was used so that a given rater would focus on inter-class variability over each scale, rather than within class variability on all scales.

Scale derived areas follows:

1. Involvement and interest of children
2. Purposeful behavior of class
3. Source of direction of academic activities
4. Nature of control over behavior
5. Effectiveness of behavioral controls
6. Quality of presentation of subject and materials
7. Differentiation of instruction
8. Teacher reaction to classroom situation
9. Reinforcement of behavior of children
10. Nature of reinforcement

With the exception of scales five and nine, there appears to be a general factor which differentiated teaching in observed classes. Intercorrelations between scales ranged between .60 and .80 and the internal consistency of the scales is well documented across all observers by scale--total score correlations of .80 and .90 with the exception of the two scales mentioned. Rater agreement on individual scales, with the exception of scale 9 varied between .50 and .60 and rater agreement on the cumulative mean rating that was made by each observer on each teacher was correlated .65.

There are important differences between raters as is reflected by their respective interscale correlation matrices. For two of the raters, the interscale correlations were generally between .40 and .60, while two of the other observers had interscale correlations between .75 and .85. Subsequent data analyses which are aimed at establishing differential effects within suburban and inner-city classes will treat observer score matrices separately in order to access the validity of different observational points of view with respect to predicting change in diverse educational settings.

Data obtained from scales was unequivical in showing suburban classes to be uniformly superior to inner-city classes. Seventy-five percent of the suburban classes were above the median and seventy percent of the inner city classes were below the median which was highly statistically significant on "t" test.

Differences between inner-city and suburban classes were statistically significant on all scales except 5, effectiveness of control; 9, reinforcement of behavior; and 10, the nature of reinforcement.

Thus, observational ratings clearly distinguish inner-city and suburban classes on selected scales and on mean rating over all scales. However, 30% of the classes overlap, five suburban classes being below the median and six inner-city classes being above the median.

These observational data will be used in order to modify the prediction of change in inner-city and suburban classes in order to determine whether high quality (as here defined) classes in inner-city schools are associated with changes in children in high quality classes in suburban schools and, similarly, whether low quality instruction in the suburbs is associated with low quality instruction in the inner-city.

DISCUSSION

This carefully structured observational survey demonstrated the degree and kind of difference that is manifest between inner-city and suburban classes. This is backed up by a film survey of selected classes, kindergarten through five, in inner-city and suburban schools. There is a close correspondence between filmed behaviors and those that are reported in the data analysis of the scales used by observers. In both cases it is apparent that inner-city schools are characterized by relatively uninvolved children, classes with extremely restricted purposes and teachers who tend to pervasively control materials and children. This control is often expressed as coercion and threats and is accompanied by a rather pedestrian presentation of materials with relatively little differentiation of instruction. Inner-city teachers appear to enjoy their teaching less than suburban teachers. These differences are quite apparent in the films, which are presently being prepared for showings at several national conventions.

Inner City and suburban classrooms will be displayed simultaneously on two adjacent screens in order to bring these comparisons into relief. Films have been subjected to detailed analyses in order to refine scaler differences. Films of the inner city and suburban classes have been combined with films of Head Start classes in order to specifically and objectively present a cross sectional longitudinal comparison of the experiences that children have in preschool, kindergarten and through the grades. The films vividly portray the contrast between selected Head Start and selected elementary school classes.

All filmed sequences have been coded according to a curricular scalematic devised by Garfunkel (1967) which identifies activities according to curricular classification (activity, substantitive or routine), substantitive or activity category (construction, performance,play gratification, language, social science, snacks, clean up or rest), process focus (mechanistic routine, skill, perceptual, cognitive or social) and control (teacher or child dominated). Each sequence is also rated on the scales developed by observers. This allows for matching of contrasting curricular and stylistic sequence across and within location (inner-city-suburban) and grade level (Head Start and Kindergarten through Grade Five). Furthermore, it provides a basis for comparing filmed sequences on ten classes to observed, recorded and rated behaviors in 38 classes which were selected by using systematic and random sampling procedures. The validity of the films is, therefore, based both on techniques and methods of selecting classes and filming then, and analytically, by obtaining comparable data on films of a limited sample of classes and anecdotal reports and ratings on a representative sample of both inner-city and suburban classes.

Preliminary findings from these studies document wide variations across Head Start inner-city and suburban classes. The obvious next step is to follow children who have been exposed to certain styles of teaching and to compare their responses to elementary schools that offer similar and contrasting classroom environments. This can serve as a control for predicting how high and low changes on various measurement procedures will respond to continuous and discontinuous learning environments. Of particular interest will be the interactions between Head Start and elementary school stylistic variations on selected measures of achievement and social-emotional behaviors.

REFERENCES

Bellack, A.A., Kliebard, H.M., Hyman, R.T., and Smith, F.L.Jr., The Language of the Classroom, New York: Teachers College Press, 1961.

Bruyn S.T., The Human Perspective in Sociology: the Methodology of Participant Observation. Englewood Cliffs, New Jersey: Prentice Hall, 1966.

Garfunkel, F., Observational Strategies for Obtaining Data on Children and Teachers in Head Start Classes, (OSOD), Boston: Boston University Headstart Evaluation and Research Center Final report to Institute for Educational Development, 1967.

Mahan, T.W. Project Concern: An Interim Report on an Educational Exploration. Hartford, Connecticut: Project Concern, 1967.

Myrdal, G. An American Dilemma. New York: Harper, 1944.

Sexton, P.C. Education and Income. New York, 1961.

Webb, E.J., Campbell, D.T., Schwartz, R.D., and Sechrest, L. Unobtrusive Measures-Nonreactive research in the Social Sciences. Chicago: Rand McNally, 1966.

Appendix A April 1, 1967

HEADSTART EVALUATION AND RESEARCH CENTER

Boston University

Scales for Rating Participant Observational Reports of Headstart Classes

1. Attitude towards teaching situation

This scale is specifically aimed at a judgement of whether the teacher enjoys the teaching situation and not whether she is a good teacher or whether the observer likes her. At the high end of this scale such adjectives as happy, pleased, exhilerated, joyful, and so forth. At the low end of the scale, unhappy, miserable, sad, pained, and so forth. The judgement revolves around what the observer sees in the behavior of the teacher and not a projection by the observer as to whether O would be happy doing the things that the teacher is doing. This, as well as other judgements, will depend upon evidence that is collected in the course of observations, and it should be possible to sight that evidence. Therefore, it is theoretically assumed that the total behavioral protocal is reducible in such a way as to provide bits of evidence to support each scaler judgement. Without such reducibility, the judgement becomes simply a "gut reaction." While admitting that the "gut reaction" is an important part of perception and judgement, the process of collecting evidence and making judgements should force the observer to look deeply into his reaction and to make essentially two judgements: the first one being whether or not he can make a rating, and the second being conditional on an affirmative response to this. The condition of being able to make the rating will always depend upon the articulation of evidence to support a given judgement.

2. Teacher's differentiation of children and activities

At the end of the scale we have a teacher who runs a class that has a high rating of individual instruction and who does not make demands upon groups of children to do the same things at the same time. High differentiation would involve either one of two strategies: a.) where there is a special plan for each child depending upon his abilities and attitudes and b.) where each child is allowed to go his own way and to seek out his own kind of activity and activity level. Low differentiation would be evident by a preponderance of classroom activities which involve all children. It does not follow from this that this scale will necessarily correlate with good teaching or poor teaching, but that it represents a style of teaching with respect to dealing with individual children of groups of children.

3. Predominent Emphasis of Curriculum

This is essentially a nominal scale which calls for a judgement on the part of the observer as to which of the categories suggest the principal manifest goals of the activity being observed. The extent to which these categories are ordinally related depends upon a presumed value system with regards to desired goals of preschool teaching. The categories to be used in this scale are taking directions, cognitive, perceptual, social emotional and a fourth category, unclear, which indicates that no single emphasis can be inferred from observed activity. The judgement, of which category a given sequence of behavior belongs to, will depend upon the behavioral priority system that operates for a given class. For example, if a given lesson or period appears to be dominated by cognitive training but if the behavior of the children cause changes in plans and redefinition of the program, then cognitive would be viewed as being a secondary goal, and the kind of activities which cognitive training give way to would be the primary designation. It is essential that we observe classes closely and long enough so that we can make inferences about what the goals, in fact, are, rather than what they are said to be. Freeplay periods might be dominated by something like the learning of routines and/or language training. Perceptual training might very well be dominated by social/emotional considerations if the behavior of the children causes the teacher to shift the emphasis for individual children, from time to time. As has been stated for the other scales, it will be necessary for observers to present evidence for manifest goals and to distinguish between the nominal categories of this scale and overall judgement of effectiveness. A good deal of work will have to be done on this scale so that it presents the observer with a series of branching scales with alternative categories, but with a theoretical connection between the different branches.

4. Purposefulness of classroom behavior

An affirmative response to this scale will depend upon clear evidence of direction and continuity. One would expect to find a considerable amount of observer disagreement over this scale because this is particularly subject to whether or not the observer is in harmony with the teacher and is able to see the underlying goals of the class as it evolves. In order to rate a teacher as being purposeful and the class as being purposeful, it will be necessary to show evidence for continuity and direction; and similarly, inorder for a teacher to be rated as being not purposeful, it will be necessary to point out discontinuity and to show many apparent shifts in direction during the course of observation.

5. Control of Materials

The question here is not so much whether it is the child or the teacher but, rather, whether the child has a say in either the gross selection of activities or materials or in their use after they are selected, or whether the teacher dominates both selection and use.

6. Communication--Responsiveness

This question is directed at the class and raises the issue of whether, whatever is going on in the class, there is great responsiveness to it on the part of the children or are they largely unresponsive or indifferent and, if anything, following through on routines rather than being responsive to activities and to the teacher. Responsiveness is indicated by a large amount of verbal and non-verbal communication, but it does not indicate that this communication is constructive or destructive or that it is good or bad.

7. Work and Play

At the high end of the scale, work and play are undifferentiated and the teacher makes little attempt to label or construct activities as being work or play, but, rather, they tend to meld together. At the low end of the scale there is a clear distinction--certain activities are presented as play activities and others are presented as work activities.

8. This scale is for a total "gut reaction" to the teacher, class, and children; and it asks the observer to indicate, without any great demand for evidence, that he thinks a given teacher is more or less effective.

All of these scales are intended to get at ordinal distinctions between a specified sample of teachers that a given observer has been assigned. All judgements are necessarily comparative, and they will depend upon what observer has seen as a part of the observational task. It is the job of the designer of the sample to make sure that each observer has a fair distribution of teacher variability in his sample and, furthermore, that this variability is not highly skewed. This means that the assignment of a sample of classes to a given observer must be proceeded by enough observation to provide evidence for gross variability within a given sample of teacher. Samples for observers should have relatively homogeneous variance.

Appendix B
October 10, 1966

PROJECT CONCERN

T. W. Mahan, Project Director

Brief Summary of Theoretical Framework

PROJECT CONCERN, although directly related to the problem of de facto segregation, is not essentially an experiment in integration; rather, it is an experiment in educational intervention designed to counteract the limited influence of urban education on the disadvantaged. Research has described the "cumulative deficit" which the child from the low socio-economic environment tends to exhibit in his school performance--a phenomenon which is dramatically accentuated among the non-white poor--and has underlined the profound task involved in reversing the trend. A review of the literature quickly communicates the impression that the problem goes beyond special teaching techniques, enriched materials, and better programming.

PROJECT CONCERN will be evaluated by measured changes in pupil behavior. Nonetheless, it is important to outline, at least in skeletal fashion, the theoretical base from which these changes are predicted. Basically, the research stems from a conviction that changes in stimuli, environment and other input data can result in changes in response or output behavior. However, it also felt that cognitive patterns for copying with formal learning situations and the affective responses which accompany these patterns have been well crystallized at the time of school entrance. This results in the use of traditional response patterns which, for the disadvantaged, are frequently ineffective for school goals. To counteract this established tendency it seems best to present the subject with an intense and pervasive experience in a radically different environment so that new responses can be provoked. This is the first stage of PROJECT CONCERN--to create some dissonance within the pupil in terms of his usual perception of himself in relation to school and to take advantage of this period of flux by reinforcing positive behaviors and attitudes.

The second aspect of the intervention model is tied to the influence of peers as a basis for the development of role fulfilling behaviors. By placing a limited number of inner city youth (about 10% of the classroom population) in a suburban classroom these same youth will be constantly in contact with models of behavior more in keeping with school values. By limiting the impact of models which reinforce the current. ineffective behavior and emphasizing the impact of different, but reasonable consistent models, it is hoped that some "shaping" of the pupils' learning styles will take place in the direction of increased academic performance.

As a catalyst to prevent too much dissonance which might create a withdrawal and/or rejection reaction, significant adult figures who share much of the child's heritage but also exhibit the desired characteristics in terms of attitudes toward school and learning are provided in the supportive team. The effectiveness of this additional factor in the change process is a focus of the research design and, hopefully, evidence will be available at the termination of the project to determine the differential impact of the learning environment as separated from the impact of adult identification figures.

In essence, PROJECT CONCERN focuses around the change in perception, already to a large extent stereotyped, which can be accomplished by a confrontation with experiences highly charged with novelty but also in a context of interpersonal support. It is predicted that changes will take place and that they will take place in the direction of the models which the suburban youth present to the bussed pupils.

EXPERIMENTAL DESIGN

PROJECT CONCERN is designed to determine the relative effectiveness of a radically different educational environment as a preventive and corrective intervention in the education of urban youth from the inner city. The theoretical rationale for the position has been discussed above, but the pragmatic aspects must be mentioned briefly here. The "vacant seat" for pupil assignment has resulted in considerable variability in the placement with some classes having only one experimental S while others have four. This in turn has created a situation which results in the experimental Ss being spread across thirty-three (33) schools while control Ss are drawn from six (6) schools. Hopefully, this diversity will have a self-cancelling effect which will underline the impact of the experimental variable - the treatment procedure. In this same regard, it is also important to stress that the Experimental Ss not receiving external supportive services are all placed in one school system (6 schools) and that generalizations from their performance must be made with that fact clearly in mind.

Nonetheless the design seems adequate to examine the relative impact of four (4) methodologies, on the learning, attitudes, and motivations of inner city youth. These methodologies, in order of their predicted effectiveness, are as follows:

1) Placement in a suburban system with supportive team assistance.
2) Placement in a suburban system without supportive team assistance.
3) Placement in an inner city school with supportive team assistance.
4) Placement in an inner city school without supportive team assistance.

Ss assigned to treatment procedures one (1) and two (2) above are considered to be Experimental Ss since they are subject to the impact of the major variable under study: placement in a radically different educational environment. Ss assigned to treatment procedures three (3) and four (4) above are classified as controls. As described above all Ss were drawn from the same population in a random fashion. Schematically, the design is as follows:

	Experimental		Groups		Control		Groups	
Grade	With Support		Without Support		With Support		Without Support	
	N	Schools	N	Schools	N	Schools	N	Schools
Kdg.	32	8	14	3	--	--	50	1
1	38	9	5	2	12	2	40	2
2	47	9	2	2	12	2	40	2
3	30	7	7	3	12	2	40	2
4	25	6	9	4	12	2	40	2
5	41	9	6	2	--	--	40	1

The criterion variables which will serve as basis for evaluating the effect of the treatment variables (suburban school placement and supportive team assistance) can be grouped into four (4) general headings:

a) Mental Ability

   1. Wechsler Intelligence Scale for Children
   2. Primary Mental Abilities

b) Academic Achievement

   1. Reading
   2. Listening
   3. Arithmetic

c) Personal-Social Development

   1. Sociometric Status
   2. Test Anxiety
   3. Attitudes
   4. Teacher Ratings
   5. School Attendence
   6. Vocational Aspiration

d) Creativity

   1. Picture Completion
   2. Circles

These data will be collected at four points: September, 1966, as a base; May, 1967, to evaluate effects after one year; September, 1967, to assess loss during the summer; May, 1968, to evaluate effects after two years. The basic statistical tests to be used will be analyses of variance and covariance. All data will be analyzed for the interaction of the following variables with the primary variables: age, sex, grade placement, school system, and where the N permits, school.

In addition, case study materials reported on a weekly basis by teachers will be utilized in an attempt to discover patterns of growth and development. Along with this approach there will be data collected which will indicate parental involvement and attitude as well as neighborhood reaction to a child's placement in the suburbs. It is anticipated that there will be significantly greater growth for the Experimental Ss as a group, but it is also hoped that evidence as to most productive and effective intervention for pupils with differing characteristics may be revealed by careful manipulation of the results.

The techniques described above will be employed on the total samples. However, it is expected that smaller samples drawn from these samples will be used to study other areas such as speech improvement, frustration tolerance, and personality variables. The major outcomes of the Project will be evaluated from this design framework by means of the following specific hypotheses stated here as predictions. For operational purposes, a "statistically significant difference" shall be defined as a deviation of such magnitude that its likelihood of occurring by change does not exceed one in twenty.

1) Experimental Ss will have significantly greater gain scores than control Ss in:

    a) all measures of mental ability
    b) all measures of academic achievement
    c) all measures of cognitive flexibility (creativity)

2) Experimental Ss will show significantly greater decrease than control Ss in measures of:

    a) general anxiety
    b) test anxiety

3) Experimental Ss will not differ significantly from control Ss in sociometric measures of:

    a) acceptance by classroom peers
    b) acceptance by neighborhood peers

4) Analyses of teacher report data on Experimental Ss will show a pattern of sequential responses which follows the following trend for Ss who show significant gains in academic performance: uncritical acceptance by the teacher; more realistic appraisal by the teacher, but with a tendency to emphasize assets; a tendency to recall and report successes and achievements; attainment of a plateau in terms of reporting pupil behavior as being relatively unexceptional and consistent.

Appendix C

Category	Examples
INSTRUCTIONS	
A. Materials	
Characteristics and amount	teacher prepared, commercial student prepared
Content--specifically, the amount, nature, or characteristics of topics related to urban environments or problems.	
B. Lessons	
Interpretation	by teacher or student and the amount. "What would have happened if there were no Civil War."
Deviations within lessons	Does the teacher allow students to introduce or follow issues that may lead away from lessons?
Spontaneity	Does T. allow asides, immediate student reactions, etc. during lessons?
Opportunity for Participation	Does T. call on all students? Do faster ones dominate? Are slow ones encouraged and given a chance?
Individual Participation	Amount of individual reading, board work, participation.
*What are the project student's reactions during recitation? How much participation, attention, cooperation?	
C. Motivation	
Origin	teacher, children, a combination through some form of theme.
Pursuit	Does T follow children's ideas, accounts even fantasies?
Characteristics	What is discussed? How is the environment utilized?
D. Evaluation-Achievement	
Type	tests, oral statements, displays of students works. (Are project students works displayed?)

Category	Examples
TEACHER-STUDENT INTERACTION	
A. Humor	Does T utilize humor to include students as opposed to ridicule.
B. Address	How does she address individuals or the class? "Boys and girls." "Students." "Children" Last names--first names.
C. Feelings	Does she express or discuss her own feelings and attempts to elicit those of the students?
D. Reward-Punishment	How does she express her favor or disfavor. "I'm proud of you." "I like obedient children."

*Examples of specific interaction with project students.

Category	Examples
DISCIPLINE	
A. Form	
Verbal-direct	"Sit down." "Don't do that."
Verbal-indirect	"Please write the word." "Why don't you put your books away."
Auditory	Clapping the hands, striking the piano
Visual	The evil eye
Physical	Holding, touching, etc.
B. Amount	How many discipline instances during any one visit.
C. Tone	Must the class be completely silent. How much noise is allowed.
D. Consistency	Is the teacher consistent with her rules and enforcing them?
E. Student pressure	Are there occasions when students discipline or assist the teacher in this area by bringing pressure upon others. "Sssh, be quiet."
PHYSICAL ORGANIZATION OF CLASSROOM	
A. Characteristics	Straight rows, tables, clusters of two and three desks.
B. Room divisions	Are their study areas, work areas, hobby areas, reading areas, etc.
C. Interaction	Does room organization assist teacher-student and student-student interaction.
STUDENT INTERACTION	
A. Opportunity	Does the seating, lessons, and assignments allow or encourage interaction. Learning groups, work groups, teacher's assistants.
B. Characteristics	Describe interactions. Students selecting one another to write spelling words on board, or to clean the desks, etc.

*Degree of project student's "mix." Do they choose others, are they aggressive, moderate, or retiring in their interactions.

APPENDIX D

April 1, 1967

HEADSTART EVALUATION AND RESEARCH CENTER

Boston University
School of Education
Boston, Massachusetts

CLASSROOM OBSERVATION AND WRITTEN REPORTS:

INSTRUCTIONS FOR OBSERVERS

INTRODUCTION

In order for us to most effectively use your observations of classrooms, it will be necessary for us to have several kinds of reports which will reflect, in a variety of ways, the teacher and child behaviors which you have observed in the classes assigned to you. These reports must be detailed enough and must include sufficient affect so that other readers can read a series of reports and rate them in ways similar to the ways in which you will be requested to rank and rate the various classes that you are observing. This does not call for the suppression of your biases, but rather the ready admission of them and explicit attempts to distinguish between those behaviors which you take a liking to as differentiated from those behaviors which you think are of high quality. This means that you have not only to observe and report what you see, but also to assimilate what you see into the working model that is represented by your ideas, feelings, and experiences. We shall bring together the various models of the several observers into an integrated framework which is controlled partially by the outline which was distributed and, further, by a series of scales which will be presented to you after you have concluded your observations.

The process of abstracting from classroom behaviors to your observations, and then to your written reports and then, still further, to a series of relevant scales is a difficult one which will depend on the kinds and degrees of differences that are found between the various classes that you observe. Difficulty is, at the same time, a function of the differences that exist within any one class over a period of time. The process that is being constructed will give a more or less clear indication of whether classes are describably and meaningfully different and, to a lesser extent, the degree of differences between these classes. The reliability of the process will depend upon the clarity and comprehensiveness of the written reports. It is necessary both to be able to carefully describe the classes that we see as well as to make some clear statements about how equivical or unequivical the system of measurement is when it is put to a fair test. In this case the tests will include the observations of classes by different observers as well as the ratings of the classes by individuals who have not seen them, but who have access to the written reports.

OUTLINE FOR CLASSROOM OBSERVATION

This outline, which was distributed to each observer, is not to be used as a checklist or as an observational guide. Rather, it should be used in the following way; observers should read and reread it carefully so that they are quite familiar with the various categories and sub-categories that describe a more or less all inclusive listing of behavioral possibilities in classroom situations. The outline does not represent a mutually exclusive system nor does it cover the detail which would bring it so much closer to the classroom situation. Observers should be quite familiar with it, but they should not actively use it during the course of their observations. After completing process reports, they should refer back to the outline in order to sensitize them to the kinds of information they are getting and the behaviors and situations which they should attend to on future visits to the class. The outline will be referred to again when the summary report is discussed below.

PROCESS REPORTS

These should include a detailed statement of everything that is observed in the classroom including the behaviors of the teacher and children, the physical characteristics of the classroom, the materials that are used and any other observations which are pertinent to discussing the class. These reports are to be thought of as the total of the observer/class interaction and they should not exclude the observer and his feelings from the report.

Observers will differ in the way in which they construct this process report, but the end result should be pretty much the same. Some of you will take notes as you are observing the class, others should write out a detailed report immediately after you leave the class, still others might develop a system for sketching out their observations so that they can then be transcribed into a running commentary describing what was seen and how it was seen.

These process reports are the raw materials for everything that follows and a single report should be made out for every observation of the class. Therefore, each observer will have at least two and preferably three process reports on each class that they observe.

It is hoped that these reports will not simply be a rather dry chronological listing of everything that happens but that they will include appropriate adjectives and interpretations that are a part of the observational process. The total interpretation of a given teacher and classroom will come in a later report. What we are interested in here are the more minute interpretations of the specific behaviors that are observed. Although we are not specifically attending to fragmentary quantitative questions such as how many times a given child is reprimanded or how often the teacher talks opposed to how often the children talk. But we should be quite aware of duration and quantity and appropriate notes should be made about persistent kinds of behaviors that take place.

The process reports will be used in two ways: in the first place they will be used by independent readers who will make judgements about the classes from reading these reports; in the second place, they will be used to document the findings of this survey and relevant parts of these reports will be abstracted and integrated into a total report of all classes. In both uses of the process reports it is necessary to have writing that is provocative and comprehensive and that projects the reader into the classroom so that he gets a feeling for what is taking place and how it is taking place.

ANALYTICAL REPORTS

There should be an analytical report for each visit to a classroom. This report represents the observer's explanation and synthesis of what he has seen. It can draw uponthe material from the process report but it is not an observational report as such but rather a critical appraisal of the classroom for the period of time that was observed. If there is no substantial difference between several process reports, it is possible to combine several of these into one analytical report. However, in general, there will be a separate analytical report for each process report.

The analytical report should refer back to the outline and should assess which parts of the outline are most relevant for the class under consideration, and what kinds of information are not readily obtainable either because of the structure of the class or because of the accident of having observed a particular kind of class or a particular segment of the curricular.

SUMMARY INTERPRETATION

There will be one summary interpretation for each teacher that you observe. This will draw upon the several process reports and analytical reports and it should integrate all of the material that you have in your possession. This summary report should have two sections to it: first, an open-ended judgemental and inferential report describing the essential of the observed behavior of the period of two or three observational periods. It should be completely openended (projective) in that you are free to draw on any material that you have in any of the visits and you should underline freely as you see fit. The second part of the summary report should closely follow the outline and should comment on each of its major sections. If there are many omissions here then it should be clear that you have not observed the class either a sufficient number of times or sufficiently long enough on any one time. We continually have to address ourselves to the question of whether we have observed behaviors which make any particular class comparable to other classes.

Classroom observation is continually plagued by the lack of comparability of data. In one class a teacher may do a large amount of talking and it might be considered to be extremely important in assessing her effectiveness. Another teacher may also do a lot of talking but it might be trivial compared to other behaviors which she displays in her work with children. This means that the problem of describing and evaluating teachers has to consider more and less effective behaviors as well as behaviors which are not applicable in an assessment of effectiveness.

Somewhere along the line, we must make judgements which stem from our descriptions and which say something meaningful about the degree and kind of impact a particular teacher might have. We must obtain a sufficient amount of material on teachers to make judgements about how effective they are with respect to the teaching of academic subject matter, of providing an environment for individual self-determination, and encouraging appropriate inter-personal relationships between the teacher and the children and between the children.

APPENDIX E

May 1, 1967

HEADSTART EVALUATION AND RESEARCH CENTER
Boston University

Scales for Rating Elementary School Classes*

1. Involvement and interest of children

Indifference, Apathy		Curiosity, Absorption

2. Purposeful behavior of class

Aimless Wandering		Direct Responsive

3. Source of direction of academic activities

Teacher		Child

4. Nature of control over behavior (Teacher)

Coercion, Threat		Trust, Respect

5. Effectiveness of behavioral controls (Teacher)

None, Class out of control		Complete, Class well controlled

6. Quality of presentation of subject or materials (Teacher)

Pedestrian, Routine		Creativity Variety, Innovation

7. Differentation of instruction (Teacher)

Monolithic, Uniformity		Highly differentiated, Individually discriminated

8. Teacher reaction to classroom situation (Teacher)

Unhappy Hostile	Indifferent	Happy, Involvement with children, Obvious enjoyment

9. Reinforcement of behavior of children (Teacher)

Not apparent		Frequent

10. Nature of reinforcement (Teacher)

Negative, Punitive, Threatening	Bribery	Positive, Approval, Encouragement

*All teachers observed by a given rater are to be sorted into five categories so that two-thirds of the teachers are in categories 2,3, and 4; one-third are to be 1 and 5. Category 1 is the left hand side of each scale and category 1 to the right hand side. Category 3 is an intermediate category.

APPENDIX F

VARIABLES FOR OBSERVATIONAL SCHEDULES (WITH SELECTED REFERENCES)

Variable Types

Object	Independent Characteristics	Dependent	
		Behavioral	Curricular
Pupil	School background Placement procedure, 2 Diagnostic Information: 2, Aptitude Achievement, Personality Family-Home	Problem solving, 2 Motivation, 2 Attention, 1 Curiosity, 1 Activity, 2 Origination Mobility Participation Disruption Individuality, 2 Pupil-Pupil Interaction, 2 Sociometric variables	Grouping, 2 Getting help, 1 Independent Activitity, 1,2
Teacher	Education Experience Age Sex Certification(s) Professional Organizations and Journals Attitudes, 3	Preparation, 2 Direction, 2 Presentation Variety,1 Sequence, 2 Verbal-Nonverbal Management-Discipline Empathy-Support-Humor,1 Evaluation-Criticism, 2 Reinforcement-Rewards, 1	Use of curriculum guide, 2 Textbooks, workbooks, 2 Teacher-prepared materials, 2 Evaluation-Reports, 2
Pupil-Teacher Interaction	Not applicable	Direction-Initiative, 1 Social Organization-Teacher or pupil centered, 1 Delegation of responsibility	Differentiation, 1,2 As related to content and procedure
Classroom	Demographic Location-Type of community Size Equipment Supervision-reported Level	Supervision-observed Climate, 1 Routines, 2 Discussion, 2 Competition, 2 Order-Disorder, 2	Content, 1: Academic-Vocational-Crafts-Social-physical and recreational Subject or project Consultants-music, art, physical education, 2

1. Classroom Observation Code Digest (Cornell, Lindrall, Sarpe, 1952)
2. Schedule for observing special class for mentally retarded children (Blatt, 1963)
3. Minnesota Teacher Attitude Inventory (Cook, Leeds, and Collis, 1951)

September 1, 1967

Frank Garfunkel
Boston University.

Head Start Evaluation and Research Center
Boston University

OBSERVATIONAL STRATEGIES FOR OBTAINING DATA ON CHILDREN AND TEACHERS IN HEAD START CLASSES (OSOD)

Curriculum

Classroom organization

Teacher style

Child reactions

Communication patterns

Acknowledgement

These instruments were developed with the active cooperation of Dr. Carolyn Stern, Director, U.C.L.A. Head Start Evaluation and Research Center and her staff, and drew heavily on instruments that had been developed, or partially developed in Head Start E. & R. Centers at U.C.L.A., Banks Street College of Education, University of Texas and Boston University.

TABLE OF CONTENTS

INTRODUCTION 2

STRATEGIES 4

Instructions for recording
    CURRICULUM PROTOCOL
        Class 6
        Child 11
    BEHAVIOURAL CATEGORIES 12
    TIME VARIABLES 13
    TEACHER-BEHAVIOURAL STYLE 13

CURRICULUM PROTOCOL 15

Key for CURRICULUM PROTOCOL 16

BEHAVIOURAL PROTOCOL 19

TIME VARIABLES PROTOCOL 19

TEACHER BEHAVIOURAL STYLE PROTOCOL 20

BEHAVIOURAL CATEGORY SCALES 21

TIME VARIABLES SCALES 22

Time allotments and sample size for teacher, class and curriculum observation
    CURRICULUM PROTOCOL 23
    Individual child observation 23

Exploratory observation and training 26

Selection of observers 27

Analysis of observational data 27

INTRODUCTION

The instruments to be described herein have been developed to obtain relevant data about the curriculum, classroom, teacher and peer effects on individual Head Start children. The techniques were developed to study seemingly important sources of variation. Since the latter dominated the former, there are a variety of methods employed, each of which would appear to be most suitable for studying a corresponding parameter. Therefore, methodologies vary according to the nature of the variable being measured.

No single strategy will satisfy the diversity of situations, teachers, methods, classroom organizations, adult interactions and activities to be found in Head Start and other levels and kinds of classrooms. Following around teachers might be appropriate in some cases but, too often, the teacher might have two rooms, administrative duties, or perform special functions in the classroom which would not lead to comparability. Observing the whole class prevents attention to important details of individual interactions between children and between teachers and children. Selecting the seemingly important incidents that come to the attention of observers is fraught with several kinds of biases - those due to importance and selectivity. Furthermore, it has become more and more apparent that we should be directly attending to the impact of adults in the classroom as they come in contact with children. On the other hand, we cannot ignore the organization and substance of the class.

Therefore, a dual strategy has been developed. There is, admittedly, a gap. The curriculum protocol, when used with the total class, can only go so far and it has been pushed to its limit. Following individual children around can be disconcerting to the observer because of a concern for behaviours that are observed but cannot be recorded.

At any rate, unbiased sampling, over time, should give best estimates for individual children (so that this might be considered as a factor in the change of individual children over time) and also, the ratings and times can be accumulated over children in a given class so that we can have central tendency measures for the class with respect to the accumulation of individual behaviours of children vis-a-vis teachers and other children.

The important thing to see, in order to understand the power of the instrument, are the ways in which two, three and four scales can be interrelated in order to get at more complex effects.

The most critical scales (process focus and control) are crucial to both strategies so they must be clearly understood by the observer. They will, of course, be controlled by other scales (participation which will be mediated by child behaviour and teacher style).

The strategy does not preclude the use of anecdotal reports. Once an observer has mastered the system (gains fluency) recording will be done in a matter of seconds and there will be more than enough time to take notes on individuals and classroom situations. Readers can fill out protocols from the anecdotal report and disparities can be studied.

Validity

We have several questions: Do the scales have face (or construct) validity? Do they include important variables that do make a difference? Is the content sufficient to differentiate quality from mediocrity? Are there any important ommissions which would prevent the scales from giving a valid estimate of potential impact?

This is being done by having individuals with diverse views review the content and discuss it. It also can be done conjecturely by comparing the content of the scales (in terms of second order interactions - taking two and three variables at a time) to hypothetically superior and inferior classes, including those that may not fit any given model in all respects.

Empirical validation will not be possible in any short period of time. However, an empirical test of the internal validity of the scales is necessary if they are to be used on any large scale. This involves using multiple observers in at least twenty classes.

Strategies

There are two tasks to be done in order to complete these scales:

1. Observer (O) spends one, two or three hours in classroom, on two, three, or four separate days, attending to total functions of class. Two entries are made every five minutes, which approximates two and one half minutes per entry, but it leaves the observer some flexibility. All teachers, aides, volunteers and student teachers are observed working (or not working) with all children. The aim of this is to get at classroom organization and curriculum. In addition to the time dimension, the number of children in any given activity is a basic dimension.

2. Observer spends one, two or three hours in classroom and attends to one child at a time for periods of ten minutes. Each child is observed on two, three or four separate days. Entries are made every thirty seconds; at the end of ten minutes seven judgments are made regarding selected teacher characteristics. The thirty-second entries include category ratings (5) and the recording of the presence of three behaviours.

Protocols and coding sheets are notated with letters and subscripts so that it will make it easy to refer to them. An index of protocols and keys follows (see pp. 14-21):

(2)

A. Curriculum Protocol for 20 minute observation of class. Three of these will be needed for each hour of observation. This is used for recording activities of total class and also, but by a different observer, for recording activities of individual child when he is observed for ten minute segments.

B. Key for scoring A when used either for toal class, or for individual child. B consists of three curricular classifications which are subdivided into five scales. This is meant to be an all-inclusive and mutually exclusive category system, i.e., everything that can occur in a preschool classroom is included, and included only once.

C1 Protocol for category entries for observation of individual child on five scales directed variously at order of activity, teacher-child relation, child and teacher. Entries are to be made once every thirty seconds on all five scales. Child ratings will be made on a previously designated child. Teacher ratings will be made on teacher who is attending to child at time of observation. If several adults interact with child, then they are all included at the appropriate time. This is not meant to be a record of any one teacher's or aide's behaviour, but rather of the behaviours of any adults that come in contact with the child. This then becomes a sample of the behaviours of all teachers with all children.

C2 This protocol includes four scales which ask for the recording of the presence of teacher-child and child-child attention and communication interactions within fifteen second intervals. A distinction is made between attention and verbal communication.

C3 Protocol for rating interaction of adult with child during ten minute period. If more than one adult interacts with child, rate the adult who has attended to him the longest. If several adults have interacted equally, take the first one. Checks can be entered in appropriate cells whenever given behaviour occurs or does not occur, and the summary rating can be made at the end of ten minute interval. Ratings are made after every ten minute observation of child.

D1 Coding sheet for C1. Includes categories within scales. All scales range between 1 and 5 except scale number 5, which includes an 0 entry for the situation when no teacher is attending to or ignoring the child in a given time interval.

D2 Description of scales in C2.

Except for C3, all entries are made consecutively at prescribed intervals, throughout stipulated time periods. This means that observers will have to be quite fluent with scales, categories, notation and the range of values of each scale. Once this fluency is obtained, the observer will be able to fully attend to class, teachers and children or child, without continually referring back to coding sheets. Movies can provide an excellent way of gaining fluency, except for the classroom use of the curriculum protocol. A library of ten-minute sequences on individual children has been developed in order to train observers and obtain rater agreement data.

INSTRUCTION FOR MAKING ENTRIES INTO CURRICULUM PROTOCOL FOR 20 MINUTE OBSERVATIONS OF CLASSES (see pp. 15-18)

Five unit entries are to be made at approximately every two and one half minutes, or two entries each five minutes. Activities that take under two minutes are excluded. Periods of rapid change are recorded partially by putting X's into unscoreable categories. For example, if a child switches activities for four or five minutes, entries are made where they reflect behaviour that lasts over two minutes. X's are entered where rapid changes do not permit a single entry for the interval. If a child is switching on all categories, the recording would be five X's. If he is fairly constant in any one, it would be coded and the others would be X's.

Observer continues to make entries at two and one half minute intervals throughout designated time period. Each protocol covers twenty minutes. If an observer stays for three hours in a class, he can fill out nine protocols, or, if ten minutes is taken between each protocol, he can fill out six protocols, two hours of recording, in a three hour period.

Key B is a category system to be used for making entries into the curriculum protocol (A).

Each entry will consist of five units which can be made alphabetically or numerically. The numerical will be easier for punching cards, but the alphabetic system may be easier for the observer-recorder, particularly during the training period.

A five unit entry will be made for each activity that goes on for over two minutes and less than three minutes. If an activity goes on for over three minutes, then it is recorded again with appropriate changes, including a change in scale 5, from I to C. When an activity terminates, the last entry for scale 5 will be a T.

An entry will be made in a given time interval for every separate activity going on, no matter how many children are involved. Therefore, the number of entries can be equal to the number of children if all children are engaged in separate activities. If a number of children are doing substantially the same thing, but not in a group (either completely separately or a parallel play or activity), then it is only necessary to make one five unit entry and put a circled number underneath it, indicating the number of children involved. Entries will be made in columns indicating sequential time intervals, and in rows indicating the number of children involved in an activity. Note that there is very little space allotted to Total Class because there can be only one entry there, and there is much more space allotted to Individual and Diads and Triads because there can be numerous entries and more space may be needed. Thus, in any one column there can be as few as one entry (if the total group is involved) and as many as the number of children in attendance (if each child is doing something different).

Note that there are five columns in each two and one half minute time segment. The five spaces for any given activity will be filled with five values corresponding to the five numbered columns of B. The first unit of each entry will be one of the primary categories from column 1 of B. Thus the observer will have to select a most appropriate primary category (Co or Pe or Py or Un or Se or Qu or La or So or Sn or Cu or Re) depending on what is going on for a particular individual or group. (It is not necessary to code curricular classification [column 0] as we can do that later. This column is for the logical convenience of the observer so that he can focus on the most appropriate primary because of the logical operational focus of the curricular classification. Note that primaries are numbered from 1 to 11, 1 to 3 being under Activity focus, 4 to 7 being under Substantitive focus, and 8 to 11 being under Routines. I have added an 0 under Activity focus because of the likelihood that some children will be wandering around and it will not be possible for the observer to stipulate a primary.)

For each primary, respectively, there are corresponding secondaries. For Activity focus each primary has unique secondaries. For Substantive focus (II) the secondaries cut across all primaries. Therefore, for II, each primary is rated Tx or De or Ex. (This is preferable to having unique secondaries for each primary, as it permits cross tabulation across primaries.)

Process focus, control and sequence (columns 3, 4 and 5) all cut across primary-secondary entries. Thus, no matter what the primary or secondary, a category of process (Co or So or Me or Na), and a category of control (A or M or P), and a category of *sequence* (*I* or *C* or *T*) must be selected and entered into an appropriate cell.

Now to give a few examples. If a group of three children are with a teacher constructing wooden boats, under the direction of the teacher, where the purpose is to construct the boats and teach the children to use simple tools, and this activity has been going on for over three minutes, then the coding would be: *Co/Fu/Sk/P/C*. From this entry it follows that the activity is *activity focused*, the activity being *construction*, the *secondary* designation being *functional*, the *process focus* being *skill*, the *control* being with the teacher and the *sequence* being *continuing*. The entry would be made in the second row (from the top).

If a child were wandering around picking fights with other children without supervision from a teacher, and if this activity persisted for at least two minutes, then the entry to be made when this began (for the first two minutes) would be *Un/In/So/A/I*. The entry would be made in the first row as the child is acting individually.

If the total class is listening to a story read by the teacher where the story is an end in itself, and there is no elaboration, and it is in its last two minutes, then the entry is made in the bottom row and is *Pe/Dp/Me/P/T*.

On the other hand, if there is an obvious attempt on the part of the teacher to elaborate the story, explain vocabulary, question the children on what is happening, then the entry is *La/De/Co/P/T*.

For convenience we will refer to scales by *curricular classification* number (a Roman numeral) and the column number. Thus, *II-3* is the *process focus* scale under *substantitive* focus.

All scales under column *1* are strictly nominal - there is no implied ordering. Scales under column *2* are nominal except for *II-2*, where there is an ordinal hierarchy. The value recorded (or the coded alphabetical symbol) should be the "highest" observed during the time interval. There is an implied order in III-2, elaborative being a higher level than functional. Again, the highest level is dominant, and if it takes place during the interval it should be so coded.

Process focus scale is partially ordinal, in the following order, from lowest to highest, in parenthetic clusters: (Me); (Sk, Pe); (Co, So). Thus, if manipulative skill training is present, but there is also a noticeable amount of attention to co-operation and sharing, then the higher level (So) would be entered. In order for the higher level activity to be entered, it must be more than a passing remark, question or answer. It has to be sustained and children must be responsive.

Control scale (column 4) is clearly ordered. Assigned value should reflect dominant influence on activity, not simply presence or absence.

Sequence (column 5) is a simple sequential scale.

The bottom row of the protocol has space for entering attendance during each interval. If children stagger in during the first hour, this will change during each time interval. Attendance includes all children present during the given interval, no matter what they are doing or where they are doing it, just so long as they are a part of the class.

In the row labelled Adults, record number of adults with class during each time interval.

The principal problem of this procedure will be for highly differentiated classes where many children are functioning individually or in diads and triads. This will be particularly troublesome when there is more than one room or when a lot of the activity takes place in a large outside area. IN SOME CASES WE MAY NEED TWO OBSERVERS, dividing up the area, so that they can give justice to individual and small group activity.

INSTRUCTIONS FOR MAKING ENTRIES INTO CURRICULUM PROTOCOL FOR TEN MINUTE OBSERVATIONS OF INDIVIDUAL CHILD USING CURRICULUM PROTOCOL (A) (see pp. 15-18)

The same protocol as is used to describe the curriculum of the class is used to describe the curricular activities of a child at the same time that scales C1, C2 and C3 are being used. Rather than asking what is going on in the class, this calls for asking what kind of activity, with whom, and when, is the individual child involved. Thus, if the child is engaged in an activity with four other children, it will be entered in the third row (from the top). When using the protocol in this way, there will only be one entry per column (or time interval) as the child cannot be in a group and functioning individually at the same time. However, the organization of the class into group or individual activities can be noted with checks in appropriate cells.

The observer can elaborate on the protocol, underneath the entry, in order to provide more detail.

As with the classroom use of this protocol, the number of children in attendance and the number of adults working with children should be recorded in each time interval.

The time interval of ten minutes is an educated guess. This means that we can observe approximately four or five children per hour, leaving some time for switching children, note taking and completing ratings. If observer stays the whole three hour period, he can get three ten minute recordings on each of four or five children. We can vary the length of time interval, the number of times we observe a given child in a day (and consequently the number of children that are observed in a day). The important point here is that whatever the interval, we should always get two or three separate time intervals on each child selected in a given day.

INSTRUCTIONS FOR RECORDING BEHAVIOURAL CATEGORIES FOR TEN MINUTE OBSERVATIONS OF INDIVIDUAL CHILDREN (C1 and D1) (see pp. 19, 21)

Five scales, each with five categories (with one exception, 5) are to be used in making entries every thirty seconds for ten minute periods. Again, the selection of thirty second intervals is an educated guess. We will see how it works and reduce it or expand it accordingly. All scales are ordinal with categories that are not meant to be all inclusive, but are mutually exclusive.

The control scale is directed to the source of initiating and terminating activity, and also to the dynamics of the continuing activity (the sequence scale, 1). Essentially, it asks the question of choice. If the child is in the midst of an activity and is being directed, corrected and helped by the teacher, then the recorded value would be 1 or 2. If the child is a part of the decision process, or if he is allowed to proceed without too much assistance, then the value would be 4 or 5. A + or a - sign in the lower right hand triangle of each cell will indicate the observer's judgment of the appropriateness of the control.

The participation scale (3) focuses on the involvement of the child in whatever he is doing, whether it be desirable or undesirable behaviour. This scale asks questions of both intensity and direction. Therefore, it is clear that "activity" must be defined during each time interval. For a very active child it may be, at times, difficult to determine which end of the scale to place him. The question is whether a child can reject an activity that he has initiated and which he is continuing to control.

In order to get out of this bind, we have included Behaviour of child (4), which requires no activity antecedent, but is simply a question of intensity.

Teacher style (5) refers to the style with which the relevant adult (teacher) controls and reinforces (or inhibits) behaviour, or at least attempts to. There is an important distinction between 0 and 3. In 3 the child is being attended to, but there is little or no apparent response. 0 implies that the child is outside of the teacher's field of attention (and possibly vision).

INSTRUCTIONS FOR RECORDING PROTOCOL FOR SELECTED TIME VARIABLES (C2) (see pp. 19, 22)

Entries in scales 6, 7 and 8 are straightforward recordings of the presence and direction of teacher-child and child-child interactions that take place in fifteen-second intervals. Scales 6, 7, 8 and 9 are related in that attention is generic to verbal communication. Both can take place when the child is functioning individually (with regards to peers) or in a group. When the teacher talks to the group, without reference to the child, and without specific response by the child, it is scored 0. Similarly, attention must involve the child, in particular, or else it is scored 0.

Scales 8 and 9, communication child-child and attention child-child, refers to conversation and non-verbal communication between child and one or more peers, where there is mutual responsiveness.

INSTRUCTIONS FOR RECORDING PROTOCOL FOR RATING TEACHER BEHAVIOURAL STYLE (C3) (see p. 20)

A rating on each of seven characteristics is to be made at the conclusion of each ten-minute observation of an individual child. The "teacher" here is the teacher, aide, volunteer or student teacher, who has had most contact with the child, either while the child is in a group situation, or when he is alone with the teacher. If several adults have contact with the child, rate the adult who has the longest contact in the ten minute interval. If several adults have equal time, then rate only the first one. If no adult has had any contact with the child, or if the contact has been superficial and of extremely short duration (less than ten seconds), then score it as not applicable (0).

As a matter of procedure, the observer should make checks under 1, 2 or 3 in each of the categories whenever there is contact, according to whether the contact is an example of the category behaviour, or whether the behaviour is not apparent.

The purpose of these scales is to get at behaviours that are obviously important, but which are not easily categorized and recorded. We are primarily asking the question of presence or absence, but we have inserted an intermediary point because of the likelihood that there will be ambiguity. Values of 3 and 1 are clear indications that the behaviour has been shown or not. The behaviour does not have to be sustained over the ten minute interval, but it has to be clearly present, and a part of the teacher's behaviour, vis-a-vis the children or child, in order to be scored. These scales can be just as readily scored if the child is in a group for the interval as if he is in a one to one situation with the teacher for all or part of the interval.

Since potentially each scale can be represented by a bipolar continuous scale, each rating will consist of two poles plus an intermediary value. For example, when rating humour, the first question is whether the teacher displays humour in her dealings with the child or the group that the child is in, for all or part of the ten minutes. If the answer is no, then the teacher is rated 1. If yes, then the question is asked as to whether it is occasional and somewhat ambiguous, in which case the rating is 2. If the display of humour through facial expressions, words or laughing is obvious (although not necessarily continuous), that is, if it takes place in more than a fleeting instant, it is recorded as 3. If the teacher has no opportunity to display humour or not to display humour, then the rating is 0, which stands for not applicable.

Curriculum Protocol for 20 minute observation of Class (or individual child in class) (A)

Date ____________________ Head Teacher _________________________

Name of Center _________________________ Child _________________________________

	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
No. of children in activity																																								
Individuals (1)																																								
Groups Diads & Triads (2)																																								
Groups of 25-50% of Class (3)																																								
Groups of 50-75% of Class (4)																																								
Total Class (5)																																								
Attend (6)																																								
Adults (7)																																								
Time (8)										5										10										15										20

B

Key for scoring classroom and child Curriculum Protocol (B)

0	1	2	3	4	5
Curricular Classification (Not to be coded)	Substantive or activity Category		Process Focus	Control (with regard to child)	Sequence
	Primary	Secondary			
I Activity focus (Ac) (Focused on activity rather than substantive content)	Co: Construction (01) (Wood, paper, clay, blocks, painting, collage, cooking, etc.)	Ae: Aesthetic (1) Fu: Functional (2)	Me: Mechanistic routine (1) Sk: Skill Manipulative (2)	P: Passive Teacher dominated (1)	I: Initial (1)
	Pe: Performing (or being performed to) (02) (Music, role playing, puppets, television, house-keeping, games, etc.)	Mu: Music (1) Dp: Dramatic play (2) Ga: Games (3)	Pe: Perceptual (3) Co: Cognitive Language (4)	M: Mediated Control between child and Teacher (2)	C: Continuing (2)
	Py: Play (03) (Water, sand, clay. Swings, slides, jungle-jims)	Sm: Small Muscle (1) Lm: Large Muscle (2) Cx: Complex (3)	So: Social emotional (5)	A: Active Child dominated (3)	T: Terminal (3)
	Un: Undefined (Wandering) (00)	Ex: Exploring (1) In: Interacting (2) No: No definition (0)	Na: Not applicable (0)		

0 - Not defined, not applicable.

B

Key for scoring classroom and child Curriculum Protocol (B contd.)

0	1	2	3	4	5
Curricular Classification (Not to be coded)	Substantive or activity Category		Process Focus	Control (with regard to child)	Sequence
	Primary	Secondary			
II Substantive (Su) (Focus on substantive content rather than activity)	Sc: Science (04) (Biology, physics, chemistry, botany, zoology)	Tx: Textual (1) (Labelling, memorizing, discriminations)	Me: Mechanistic, routine (1)	P: Passive Teacher dominated (1)	I: Initial (1)
			Sk: Skill Manipulative (2)		
	Qu: Quantitative (05) (Numbers, sizes, shapes, puzzles)	De: Demonstration (2) (experimental, implications)	Pe: Perceptual (3)	M: Mediated Control between child and Teacher (2)	C: Continuing (2)
	La: Language, Verbal (06) (Speech, discussion, stories, writing, reading)		Co: Cognitive, Language (4)		
	So: Social relations (06) (Interpersonal relations, culture, social organization, history, government, community)	Ex: Experimental (3) (Direct involvement, transformations)	So: Social emotional (5)	A: Active Child dominated (3)	T: Terminal (3)
			Na: Not applicable (0)		

0 - Not defined, not applicable.

B

Key for scoring classroom and child Curriculum Protocol (B contd.)

0	1	2	3	4	5
Curricular Classification (Not to be coded)	Substantive or activity Category		Process Focus	Control (with regard to child)	Sequence
	Primary	Secondary			
III Routines (Ro)	Sn: Snacks (08) (Juice, lunch)	Fu: Functional (1)	Me: Mechanistic routine (1)	P: Passive Teacher dominated (1)	I: Initial (1)
	Cu: Clean Up (09)		Sk: Skill Manipulative (3)		
	Re: Rest (10)		Pe: Perceptual (3)	M: Mediated Control between child and Teacher (2)	C: Continuing (2)
	O: Other (11) (Arrival, departure, toileting, washing, dressing)	El: Elaborative (2)	Co: Cognitive Language (4)		
			So: Social emotional (5)	A: Active Child dominated (3)	T: Terminal (3)
			Na: Not applicable (0)		

0 - Not defined, not applicable.

Behavioural and Time Variables Protocol

| | | | | | | | | | | | | | | | |
Mo. day yr. cl. ch. obs fm

Name of Child ______________________ Head Teacher ______________________

Name of Teacher ______________________ Name of Center ______________________

Time of day ______________ Name of Observer ______________________ Matched yes ______ no ______

		Category	10 minutes - 15 and 30 second intervals 1	2	3	4	5	6	7	8	9	10
C1	1	Sequence (I-C-T)										
	2	Control (1-5) T-Ch										
	3	Participation (1-5) Ch										
	4	Behavior (1-5) Ch										
	5	Style (0-5) T										
C2	6	Attention: T-Ch Direction										
	7	Communication: T-Ch Direction										
	8	Attention: Ch-Ch Direction										
	9	Communication: Ch-Ch Direction										

C3

Date ______________
Mo. Day Yr.

Name of Child ______________ Head Teacher ______________

Name of Teacher ______________ Name of H.S. Center ______________

		Rating			
	Category	Not applicable (0)	Rarely or never (1)	Occasional, Intermittent (2)	Obvious and sustained (3)
1	Display of humour				
2	Warmth towards child(ren)				
3	Resourceful presentation of material with/without planning				
4	Spontaneity of presentation, on-the-spot flexibility in use of materials & language				
5	Direct encouragement of verbal responses in child(ren)				
6	Behaviour indicative of being happy with and enjoying teaching situation				
7	Shows respect for children by maintaining single standard in language and givingreasons for actions				

Teacher Protocol for rating Teacher Behavioural Style during 10 minute observation of individual child - rate only the adult who has contact with observed child. If several adults have equal contact, rate only the first one. Not applicable refers to no contact. (To be administered for each ten minute observation of individual child.)

(C3)

Category Scales for Time Sampling of Individual Child Classroom Behaviour (D1)

1. Sequence of behaviour: I - initial; C - continuing; T - terminal

2. Control of activity

Teacher Direction and insistence	Teacher Suggestion	Teacher Presents Alternatives	Teacher Child Collaboration	Child Controlled
1	2	3	4	5

3. Participation of child in activity

Rejection	Reluctance	Indifference	Moderate Involvement	Intense Involvement, absorption
1	2	3	4	5

4. Behaviour of child

Passive Non-responsive	Compliant Participation	Active Exploratory	Verbally Aggressive	Physically Aggressive
1	2	3	4	5

5. Teacher style of approaching, directing and responding to child (T)

Not attending to activity or task in any way	Severely punitive, physical	Verbal punitive, sarcasm, intimidation, ridicule, threat, coercion	Permissive, attending to but non-committal, uninvolved	Enabling, reassurance, encouragement, constructive, criticism	Unequivocal and indiscriminate support, over indulgence
0	1	2	3	4	5

Scales for Time Sampling of Individual Child Classroom Behaviour (D2)

6. Attention: Teacher to Child (→) and Child to Teacher (←) and two-way interaction (↔)

7. Verbal communication: same as #6.

8. Attention: Child - Child

9. Verbal communication: Child - Child.

(Presence [✓] or absence [0] to be scored every 15 seconds.)

E & ER Observation of Classes
Time Allotments

The following chart gives an estimate of the sample:

	E	ER	Total
No. of Children:	75-100	45-60	120-160
No. of Classes:	5-10	4-8	9-18
Mean No. of Classes	8	6	14
Mean No. of Children/ Class	10	10	140

Time allotment: 30 hours/class with two observers or
60 observational hours/class equals total of 840.

Time Allotment for Curriculum Period

Three 3 hour observations/class equals 9 observation hours/class (assuming 14 classes) 9 x 14 equals	126
Extra observers for highly differentiated classes	30
	156

Time Allotment for Individual Child Observation

3 hours/child overall (140 children) (756)	
2 hours/child overall	504
1 hour/child overall (252)	
Training and exploratory observation	180
	840

Observation of individual children will take more days than observation of class. Suggested guideline: Each child should be observed on at least one cycle (three 10 minute intervals in a single day) concurrently with observation of class (Curriculum Protocol); this would leave three cycles which would take place without concurrent administration of curriculum protocol.

Time for Individual Child Observation

Options on total time, assuming a 3 hour observation period, will net two hours of protocol time per day (four 1/2 hour observations per day).

Total net time each child is observed in hours	No. of cycles	No. of 10 min. observations 3/cycle	Total hours for sample, assuming an average of 14 classes, 140 children
1	2	6	252
2*	4*	12*	504*
3	6	18	756

*Recommended

Model for each class (assuming 12 children from class in sample)

Cycles - Days

Sub Samples:		1	2	3	4	5	6
	A	4	4	4	4	4	4
	B	4	4	4	4	4	4
	C	4	4	4	4	4	4

Two cycles will give one hour of observation per child in ten minute intervals, in two separate days.

Four cycles will give two hours of observation per child.

Six cycles will give three hours of observation per child.

Each cycle will take three 3 hours per day or 9 observer hours per class.

Total time for completing is 9 X number of cycles x number of classes.

Observation of individual children will be rotated as follows: four children out of the sample in a given class will be selected for a given day's (3 hours) observation. Children will be assigned numbers and the four will be selected from the sample in the class. There are several possible procedures:

1. The four can be rotated in straightforward fashion for the three hours - 1, 2, 3, 4; 1, 2, 3, 4; 1, 2, 3, 4 - four children to be observed in an hour leaves 20 minutes per hour for finding next child, completing protocols, etc. This can roughly be spaced as follows: 10,5,10,5,10,5,10,5 equals 60 minutes.

2. The sample in each class can be randomly assigned to observation periods and observation would continue until each child had been observed for a given amount of time. Using this procedure, it would be possible for a child to be observed 0 to 12 times on a given day.

Method #1 is procedurally easier, but method #2 is theoretically more sound, as it eliminates bias due to observation dependency. Method #2 would give unbiased estimates of child-peer-teacher activity interaction; therefore, it would be more defensible to accumulate ratings for children in a given class in order to get measures of total class functioning.

It would further strengthen the possibilities for inference if days for observation were selected using a random or stratified random procedure. It would seem that since the number of classes per center will be modest, it might be reasonable to do it in this way.

Ideally the procedure would go as follows:

A sample of class days would be randomly selected from all possible class days, stratifying on beginning (Monday), middle (Tuesday, Wednesday, Thursday) and end (Friday) of week and probably ignoring vacation and party days (either eliminate them entirely or take no cognizance of them and include them as they are selected). Trip days can either be eliminated entirely or included as they come up. The latter is theoretically preferable, but might present practical problems. Data from classes that include trips as a major part of their program would be biased. The individual child scales have been constructed so that they can be used for any activity that may be going on. The decision to include or exclude trip days should be made after the explanatory observation. Other considerations might also come up at that time.

Exploratory Observation and Training

One hundred eighty hours are to be used for exploratory observation and training, to be divided up 30 and 150. Exploratory observation is for the purpose of getting preliminary data that can be used in:

1. Sample selection
2. Determination of stratification variables.
3. Determination of restrictions and exclusions.

If the sample of available classes is considerably larger than 14, it might be necessary to do preliminary observation in order to obtain the 14 classes to be used in the sample. This depends on between class variability that is desirable for the final sample of classes.

In order to set up an overall observational schedule, it will be necessary to observe each of the 14 classes on two separate occasions for approximately 45 minutes each, one observation to be in the first half of the school day, the other to be in the second half. This will provide information both from observation and interviewing the teacher on the following:

1. Overall schedule of class, intra and inter day variability, trips, expected attendance, days when observation would not be feasible, classroom organization, etc.
2. Amount of differentiation (as opposed to total or small group) in class.
3. Feasibility of having more than one observer in the room at a time. (Probably the more highly differentiated classes, which might require more than one observer for the curriculum protocol, will be the classes that can most reasonably be expected to be able to have more observers without interrupting the program).

It is estimated that this will involve a net of two hours which can be done by two observers in a week as follows:

		M	T	W	Thurs.	Fri.
Observers	A	1,2,3	4,5,6	7,8,9	10,11,12	13,14
	B	7,8,9	1,2,3	13,14	4,5,6	10,11,12

Class numbers are inserted in cells.

Including travel time, this will gross out at 30 hours.

Alternatively, each observer can visit seven classes, each twice, which will give the same net and gross hours, but which may be more practical, because on the second visit, the observer will already be familiar with the setting. The first procedure has the advantage of observation on each class by different observers.

This leaves 150 hours for training observers.

Selection of Observers

It is expected that individuals with experience and training in the education of pre-school children will be included in the team of observers. This is not to say that it is necessary (or even desirable) that all observers have such training and experience, but that pre-school education and experience will be represented.

Utilizing films, all observers can be screened on their ability to record accurately curricular and behavioural sequences. Although much has been done in the development of this observational strategy to reduce observer effects, it will be necessary to continually quality control the process by using films and two-way observational set-ups where many observers can rate the same activity.

Analysis of Data

Characteristic curves of classes and clusters of classes in terms of differentiation (vertical axis of *curriculum protocol*, *A*) X time (horizontal axis of *curriculum protocol*), will be mapped. These curves can be compared with reference to *control* (*B*-*4* and *D1*-*2*), *process focus* (*B*-*3*), *participation* (*D1*-*3*) and *teacher style* (*D1*-*5*) across *curricular classifications* (*B*-*0*) and *primary* and *secondary substantitive or activity* categories (*B*-*1* and *B*-*2*).

By utilizing curve fitting methodologies and, possibly, multiple discriminant function analysis, maximally different functions can be calculated and described in order to set up hypotheses about goals that can be inferred from observational data - that is, goals that are assumed to exist because of what goes on in the class. In more straightforward terms, this analysis will develop specific predictions for single children, classes and clusters of classes regarding probabl effects of various kinds of programs on children. The first question to be attended to is, if a given class (or cluster of classes) is affecting children, what kind and degree of effects are most likely? It may be that, for some of the clusters, we will not have measurements suitable for obtaining data on behaviours that appear to have the most possibility of changing. It will be possible to make probability estimates that any given class, when compared to any other class, whill show child changes in particular areas. The categories used for the curriculum protocol (B) will provide a basis for connecting classroom procedures with individual child behaviours. Control (D1-2), participation (D1-3) and teacher style (D1-5) will provide data on how curriculum is presented and how it is received, thus allowing for inferences about not only what is being presented to the child, but how it is expressed and how it is received.

Observational recordings can be used for obtaining individual child measures as well as class measures. When a prediction about change is made for a given class, it can be mediated by the data obtained on individuals in that class. Thus, we will not have to rely on global predictions exclusively. The hypothesis that a child who has been exposed to a considerable amount of a particular kind of cognitive content will change cognitively will be mediated (different hypotheses for different children or clusters of children) by variables of the behaviour protocols (C1 and C2).

Time variables (C-2 and D-2) provide quantitative data on an aspect of participation (C1-3 and D1-3), and also an aspect of curriculum (B-3-4 and BII-1-6).

For any given curriculum, the time variables will provide verification or rejection of the observed goals, in terms of particular children.

The teacher protocol (C3) consists of seven scales, the data on which will be analyzed separately, in order to isolate two or three factors. Resulting factor scores on teachers will be used to further mediate the evaluation of attention as it applies to children. We have already discussed the questions of how, both for the teacher and for the child. For given how-how interactions, we ask whether there is corresponding systematic variation on certain aspects of judged teacher style. The aim of these analyses is to explore the question: what kind of teacher (or teaching) will get what kind of attention? The "kind of teacher" (or "teaching") is gotten at by the teacher protocol (C3) and also by D1-2, D1-5 and, more generally, from the curriculum protocol (A and B). "Kind of attention" is explored in participation (D1-3), behaviour (D1-4) and the time variables (C2 and D2).

EMOTIONAL DISTURBANCE:
PRESCHOOL CHILDREN

HEAD START EVALUATION AND RESEARCH CENTER

PRIMARY AND SECONDARY PREVENTION STUDYING CLINICAL PROCESS AND DISTURBANCE WITH PRESCHOOL CHILDREN[1]

Caroline Fish
Frank Garfunkel
Pierre Johannet
Ann Coolidge

Boston University

ABSTRACT

During the summer of 1967, a six week experimental clinical survey with a selected small number of preschool children ranging in age from 3 to 6, from two differing social class communities, evolved a methodological procedure that would assist in clarifying the issue of emotional disturbance and potential emotional disturbance in young children. The primary thrust was to set up closely coordinated interdisciplinary teams that could interact with preschool systems and their representatives on a consistent longitudinal basis. Class differences in the system and attitudes of their personnel were more striking than the actual differences between the children. Techniques of testing, observing, coordinating and recording data were established. Full-year projects based on this model have evolved in the inner city Headstart programs.

---

[1] "The research reported herein was performed pursuant to a contract with the Office of Economic Opportunity, Executive Office of the President, Washington, D.C., 20506. The opinions expressed herein are those of the author and should not be construed as representing the opinions or policy of any agency of the United States Government."

PRIMARY AND SECONDARY PREVENTION STUDYING CLINICAL PROCESS AND DISTURBANCE WITH PRESCHOOL CHILDREN[1]

Caroline Fish
Frank Garfunkel
Pierre Johannet
Ann Coolidge

Boston University

INTRODUCTION

In the summer of 1967, an interdisciplinary team addressed itself to the study of disturbed preschool children, comparing a selected sample of low and middle class communities with a view to finding a broader, more descriminating definition of emotional disturbance between the ages of three and six than is currently found in the literature. A need to study the epidemiology of emotional disturbance, especially the influences of social class, is evident. Not only were diagnostic procedures for existing emotional disturbance considered, but longitudinal development of the child was studied with the hope of predicting disturbance that might lead to school failure. In summary, emotional disturbance and potential emotional disturbance were investigated. No attempt was made to review a large population. The groups studied were a small group of Headstart children from an inner city environment and a comparable group from a middle class suburban pre-primary program. The question was raised as to finding discriminating data necessary for making useful prediction of the child's functioning in the preschool school setting. Data were studied for developing interventions to use within the classroom for the benefit of the child. The communication process of adults dealing with the child was also studied.

In particular, the process of interdisciplinary communication was scrutinized in this clinical survey. Administrators from the University and community agencies worked closely with teachers, aides, psychiatrists, psychologists, social workers, speech specialists, parents and students, sharpening communication and defining and differentiating problems. One aim of the program was to help teachers identify and cope with emotional disturbance in the classroom, to know where and when to seek help and how to coordinate their efforts with families and community efforts.

---

[1] "The research reported herein was performed pursuant to a contract with the Office of Economic Opportunity, Executive Office of the President, Washington, D.C., 20506. The opinions expressed herein are those of the author and should not be construed as representing the opinions or policy of any agency of the United States Government."

The interest, too, was in sensitizing professionals to their own profession. Teachers need to be remotivated to sense a renewal of committment. This holds for psychiatrists, psychologists and social workers. We not only need to hear and understand what other professionals are saying and doing, but we need to be more aware of definiciences in our own professions. We were particularly interested in seeing how training psychiatrists in the school setting could increase their awareness of the total functions of the child, particularly in helping them to recognize the efficacy of certain teacher interventions, groupings and styles. We wanted to see how teachers would respond to the insights of psychiatrists and psychologists in mutually observing the manifest behavior of a child.

Finally, suggestions were to be made to communities to help them develop resources within their own school systems and family services to help disturbed children. The practical problems of effecting change in established systems, such as the schools, are of considerable importance at this time. The question was considered as to how communities can be motivated to change and adapt their systems to the changing demands of the society around them.

In summary, the clinical approaches of this study led to increasing our knowledge of how professionals change in changing settings. Methodologically, participant observations of interdisciplinary groups led to sensitizing professionals which in turn, hopefully, led to benefiting the child in his variable social settings.

THEORETICAL ISSUES

Many comparative studies have been done in recent years on incidence and type of psychopathology in middle and lower class samples (Hollingshead and Redlich, 1958; Dunham, 1964; Harrison, 1965; Burgess, 1964; Luchterhand and Weller, 1965; Miller and Grigg, 1966; Shoemaker, 1965; Swift, 1966). Only a few have looked at the influences of social class on childhood disturbance. (Charry, 1967; Dairs, 1948; Bennet and Gist, 1964, Deutch, 1965; Hernandez, 1963; Hollingshead, 1949; Jensen, 1966; Miller and Swanson,1960; McDermott, Harrison, Schraga and Wilson, 1965; Sears, Maccoby and Levin, 1957). The incidence of emotional disturbance in selected groups of children and the correlates of socioeconomic status with the type of illness within the population needs scrutiny. Since few such studies have been done in children, more should be attempted in the area of diagnosis of emotional disturbance.

We also need to explore the relationship of emotional disturbance to total personality function and childhood development. We need to know more about how the child's initial experiences in school situations effects his emotional development and how existing disturbances can be ameliorated in the school setting. How does the child change and learn to adapt to the demands of the school? What information can specialists in child development, particularly cognitive development and behavior, provide that would help the teacher and the system benefit the child?

A primary difficulty of such a study is in the actual definition of emotional disturbance, as it is necessary to define emotional disturbance operationally. For this study, emotional disturbance was seen as a function of the child showing deviation in relationship to expected school behavior. There seem to be two groups of children that need examination: One is the severely disturbed group where there seems to be no question in the minds of teachers or consulting professionals that the child is in need of immediate and special help, he could not get along without it; and he certainly cannot cope with a class situation without considerable assistance from the outside.

The next group is the middle group of children who show behavior that is identified as disturbing or disturbed in the classroom. Sometimes only the teacher sees aspects of behavior that she finds disturbing; sometimes the consulting observer, whether psychiatrist, psycologist or consulting teacher, finds disturbing behavior. Often both teacher and consultant are aware of deviance in a child. It is this latter group which is of particular interest to us in this project, since a dialogue between teacher and consultant on observed child behavior could result in a richer and clearer definition of pathology. Whether this deviant behavior has a good or bad prognosis for future success in schools would depend a good deal on the type of deviation that was manifest. This is the area that the total team wanted to explore. Was this child presently showing vulnerability and maladaptive approaches in responding to immediate stress situations or was this deviation going to be of much longer duration? Could it be corrected by the child or would external assistance be needed? Was it necessary to discriminate between deviant behavior caused by intrapersonal dynamics versus deviation caused by undue environmental (low social-economic status) stress? Also to be ruled out was the question of deviation due to a deficiency or maladaptation of the central nervous system.

Several goals could be explored. The following hypotheses were generated and continued to generate theoretical and methodological issues as the survey progressed: 1). By bringing children together from two differing social classes into a diagnostic classroom to be observed by an interdisciplinary team, nosology and social class differences could be viewed simultaneously. 2). The manifestations of emotional disturbance, such as the handling of aggression, sex identification and withdrawal, would differ according to class. 3). Preliminary observation in classroom settings by the professional team, followed by three sessions in the diagnostic classroom, provided an opportunity to get at co-variation between developmental versus situational malfunctioning. To study these variables separately seemed futile, since the interaction of the two is constant. 4). Service had to be provided to the community before any community support would be given to the program. 5). The attitude of school officials, from admistrators through teachers, would be more cooperative and sustained in the middle class group than in the lower class group. 6). Enough structure could be imposed on clinical and observational procedures to facilitate inferential processes on the part of clinicians, and cut down on confounding variables. 7). Certain situations would be detrimentally handled by indigenous, nontrained personnel.

CLINICAL PROCESS

Two groups of preschool children were selected in the Boston area, one from the inner city and one from a middle class neighboring suburb with a relatively stable population. The children in the inner city came from several Headstart classes, those in the suburb from a summer public school pre-primary program. The project was instigated and coordinated by the Boston University Head Start Evaluation and Research Center. Representatives from the faculty of the schools of Education and Medicine, as well as students participated in the project. The Headstart group was coordinated by the South End Neighborhood Action Program(SNAP), a delegate agency of Action for Boston Community Development (ABCD), an inner city community action programs financed by the Office of Economic Opportunity. The suburban group was coordinated by the staff of the public school system, including a school adjustment counselor, and by a local community mental health clinic.

The University provided two classrooms with a large observation room with one-way mirrors. A master teacher and an aide ran the diagnostic class for those children selected from the populations for the purposes of generating diagnostic hypotheses. The class was studied from the observation room by two child psychiatrists, one fourth year resident in child psychiatry, two psychologists, three social workers, a speech and hearing specialist and a consulting master teacher. Observing and recording was done by students in the Boston University Psycho-Educational Clinic.

The children in the Headstart programs had been chosen primarily for their financial need and the degree of stress in the family. 120 children were in 8 classes. From these children, 26 were selected for further study, but only 19 actually came into the diagnostic class, 11 males and 18 females. Eight teachers and 8 aides were assisted by 8 neighborhood workers.

Two-hundred and fifty children, in 3 schools with 9 teachers made up the population of the suburban pre-school. These children had to be 5 years old before January 1, 1968. The pre-primary program had been started in this community because the school system did not feel it could finance a kindergarten program, but since twenty to thirty per cent of the children were failing in the first grade, the schools were backing this attempt to assist children in getting ready for school. Two of the schools were old schools in a middle income area. One school was a new school with a more well-to-do population. Nine children were recommended for the diagnostic class, eight came in, four males and four females.

Preliminary meetings were held by the University professional team with social agencies and teachers involved with these various classrooms to clarify plans and procedures. Beginning with the first weeks of operation, and then regularly throughout the summer, observers, either from psychiatry, including the psychiatry resident in training, and/or education observed in the classrooms. The visits were for at least an hour. If the teacher designated any special problem, or if a child caught the clinical appraisal of the observers, the observers returned for added observation. The teachers did not have to wait for the scheduled visits of the observing teams. If a teacher felt a child needed immediate attention she could call the University and request a visit. The teams tried to get out within two days after the request of the teacher; or a later appointment was arranged when the teachers of the University staff could discuss in detail the behavior of the child and arrange details of his coming for a diagnostic session if this seemed necessary. In several cases, the children were discussed but did not come into the diagnostic class. Specific suggestions were made to the teacher on various ways of handling the child in the classroom and this seemed to suffice in helping the child adjust.

If a teacher wanted a child worked up in detail, she filled out a form with the child's name, her name, the school, the date, the comments of the visiting teams on the child, and her specific questions on the behavior of the child. In addition, the visiting teacher consultant on the team tried to take down verbatim the requests and observations of the classroom teacher. These reports were mimeographed and ready for all members of the staff when the child entered the diagnostic class. Also mimeographed for curculation were the team's observations of the classrooms and of specific children.

Scheduling of the children at the diagnostic class was arranged through the appropriate centers. Each child was to come for three days, Tuesday, Wednesday and Thursday. The diagnostic sessions were routinely observed by the team of psychologists, psychiatrists and educators as well as students in the Department of Education in an observation room, with one-way glass and sound apparatus, large enough to accommodate the numbers of people involved.
After each session, the professional team met to discuss the behavior of the children they had observed. Predictions and suggestions were made about behavior in the classroom and how to get help outside the classroom, particularly for family problems. Groups of three to six children were seen routinely. Children from both social mileius were present simultaneously. The class was operated much like a standard nursery school program in regard to routine and materials. Two rooms were available, with the large observation room inbetween. On the classroom the children first entered, was arranged in sections. One area had toy furniture resembling a kitchen area including running water; another had round tables with chairs around them; another had shelves with toys, puzzles, crayons, etc. A counter served as an area for shaving cream play, clay work and the like. In the classroom across the hall, which the children spent the latter half of their visit each day, was the large motor equipment: slides, tricyles, jungle-gym-house, rocking boat, pyramid stairs, etc. When the children came over to play in this room, individual interviews, whether for psychiatric workup or for speech and hearing evaluation, were done in the classroom so the child did not have to be removed from his familiar milieu and so that the diagnostic team could observe.

Graduate students in training, were specifically assigned to observe and record the behavior of each child, to assist the team at the final evaluation. These recorders were instructed to note relationships with adults and peers, the attitude of the child to authority, the amount of intention and the time spent on tasks, the handedness of the child, the motor coordination, including eyes, hands, small and large muscles, and to give as much verbatim vocabulary and sentence structure as possible. They were also to look for affect, to note frustration tolerance, attention span, reflexive-impulsive set, and recorded any fantasy themes that could be picked up, and the amount of curiosity, risk and exploring the child showed. Each five to ten minute intervals was noted so the sequence of the child's play and duration of his interest was recorded.

Parents were included in the observational process during the first session. They were invited and arrangements made for their coming through their respective agencies. The parents who did come seemed to enjoy watching the children and talking informally to the team. Teachers of the classrooms and administrative staffs of the agencies were also invited to come and observe the diagnostic sessions, whether they had a child from their class present or not. They usually attended the first and second days of the three day sessions. They stayed after the children had gone to discuss individual cases and to keep communicative channels between professions open.

The third day, after the class session was over, was devoted to a staff meeting summing up the cases, generating hypotheses, and working out specific plans for the future assistance to the child, the teacher and the schools.

Following the diagnostic sessions, two members of the team went to the classroom for at least one more observation of the child after he had been seen in the diagnostic session and after he had been in his own classroom situation for a longer period of time. The interest generated in the suburban schools was great enough to warrant regular weekly meetings with teachers who had brought in children as well as the other teachers to discuss classroom behavior management, diagnostic clues, reasons for referral, etc.

Preliminary to the child's arrival at the diagnostic session, and even if he were not scheduled to come to the University but was showing difficulties, each child was tested by a senior psychologist or one of her three assistant psychologists who had all had experience in testing children of this age. Each psychologist had an observer-recorder to record the behavior of the child as well as his actual test performance on the Stanford Binet which was the standardized test used. The Binet was administered in two and sometimes three sessions, depending on the ability of the child to take the test or the need of the psychologist to establish better contact with the child.

Not only were the children tested, but family interviews were attempted. A questionnaire was drawn up by the psychiatrists and psycologists outlining the areas for investigation. The interest was in exploring familial attitudes and concerns rather than trying to get accurate family histories.

Finally, the professional team, the master teacher and her aide, the consultant teacher, two child psychiatrists, two psychologists, the speech consultant and a child psychiatry resident met in a series of intensive meetings to collate the data, draw up final reports and make specific recommendations. Two copies of the final reports of each child were sent to the appropriate agencies, one copy was kept on file at the University, and special reports were written for the teacher of each child to help her handle the child in the classroom setting.

DISCUSSION

This clinical survey raised more question than it had set out to answer. The major problem was shortness of time and limited population. In spite of this, the children who were followed were seen in a coordinated, sequential fashion by representatives of different disciplines, to report on behavior in many situations affording continuity, contrast and varying kinds of stress. Of particular interest was comparing the two social classes together operating in the same classroom under the same conditions.

Class differences in pathology were not as striking as the fact that there were so many girls. This has been baffling to the team. The incidents of pathology in boys is reportedly much higher in the general population but we saw as many girls as we did boys. We questioned if there was a social class influence in this phenomenon. The teachers and their training might have effected this sex choice. Young teachers seemed to be more lenient and understanding of boys' behavior than of girls'. Acting out among boys seemed to be acceptable to the teachers. They probably saw acting out, particularly in the inner city classes, as an adaptive function. We also felt that somehow the teams in the classroom were not getting at withdrawal symptons.

Also surprising was that not as many children were referred as expected because of aggressive acting out or anger. Aggressive acting out seem to appear lesser in the summer program. We could not detect any class differences on this score. We were also interested to find that all the children we saw were toilet-trained.

Even though the time was short, in looking at the covariation between situational and longitudinal maladjustment, marked changes could be seen in many of the children's behaviors over time. Some changed dramatically during the three day sessions, enabling us to define, categorize, and make more appropriate references than if we had only seen the child on a one-short basis.

In our experience in dealing with sub-communities, such as the schools, that have not been anxious to explore areas such as emotional disturbance, we have found that it is impossible to do decent research or even get into the schools unless you are able to promise and carry through on specific service to the children and to the teachers and the families that send children feel that some real good will come to them they will not cooperate. The visiting teachers were particularly anxious to have immediate help with problem children.

As this program continues, we will need more information than we were able to get in the process we had this summer for the benefit of the schools and our own research. Within the first month, all initial observations of classrooms and teachers in observing teams would be gathered. By the second month, family reports and psychologicals would be done. Then, in the third month, the special diagnostic class and the interviews could be done. It is most important that these children continue to be followed in the classroom as the year continues. We need more longitudinal information on these children, not only in their own development, but in general growth. We need to know a great deal more about the established values of systems of these children and their families. There is an inherent danger in providing these children with what we believe is an optimum program when this program is contrary to the values and expectations of the community involved. We can see the University playing an important role in setting up a strategy--a well thoughtout and carefully planned strategy--for getting this type of information on values and existing structures and then maneuvering for any type of appropriate, approximate social change which is considered feasible by the community as well as by the University and the research groups. We must consider the realistic setting in which the child moves and will move and adapt ourselves to these needs. There is an element of tragedy in setting up classes which are destined to guaranteed failure of the child in the setting to which he will return. On the conclusion of the third month, when all information is in on the children, an appropriate prediction should be made as to the development of this child and the possibilities of his being a school failure. Tests will have provided us not only an approximation of his personality development, his cognitive development, but also of his level of achievement. The predictions should focus on success and/or failure in school. This could then be checked out later in a longitudinal research study, which is now evolving.

Of great importance to the whole project is the question of feed-back to the community. It is most important that the agencies not only get the material as quickly as possible, with recommendations and findings on each child, but that they be responsible and prompt in getting special reports to the teachers and providing any follow-up discussion that might be necessary.

There should also be a follow-up on each child who was observed and picked up for some sign of maladaptive behavior but not serious enough to warrant referral to the diagnostic classroom. These children should all be carefully watched. It is suggested that a confidential report on the whole program go to the superintendents of schools so that they have some understanding not only of what we are doing, but of the magnitude and significance of emotional disturbance in preschool children in their areas. It is also important that SNAP have some idea of the degree and amount of incidents of emotional disturbance in their children so they can provide pressure for more specific help from the community action programs.

Our hypothesis that the attitude of school officials would be more cooperative and sustained in the middle class than in the lower class group was upheld. One of the major thrusts of this program was to help teachers understand emotional disturbance. It was of particular interest that the teachers in the suburb were more interested in the experiment came voluntarily more often, stayed longer, including the teachers who did not have children going through the diagnostic process, than the teachers in the inner city. The staff expressed more sympathetic rapport with the teachers in the suburb than they did with the teacher in the inner city. This was probably due to the fact that the city teachers were hired on a temporary basis, were not part of an established system, would be leaving the community at the end of the summer. Finally the whole general program in the public schools was in abeyance at the time of this program. Therefore, there was a certain hopelessness for all of us in approaching these city children, in trying to decide how effective, if effective at all, we could be with them, whereas the suburban system was well established and had an aura of performance. Besides the staff was excited by the newness and the creative approach of the University program.

Several suggestions were made as to how greater effectiveness could be achieved with the teachers in the over-all program. In the first place, it was suggested that the teachers should have written longer, more specific reports on each child they referred. These reports should have been followed up in detail by the psychiatrists and discussed with the teachers as to the degree of importance of their observations. Also suggested was that teachers go around to all classrooms to see how other teachers considered and dealt with children thought to be disturbed. The teachers should talk more with each other and with consultants. As a first step, every teacher could be invited in to observe the master teacher teaching a class of emotionally disturbed children. The second step could be to have a teacher choose a fellow teacher and the two of them go in compatible combinations to visit and attend other classes. The students in training at the University would be able to takeover the teacher's role in the classroom which would give the student experience and release the teacher for time to observe. Teachers' aides could also help in taking over the class so the teacher could be released. Specific seminars for the teachers could be used to discuss appropriate ways of getting help for their own children, how to observe, how to make choices on interventions, how to elicit active, immediate intervention from the administration. Primarily, the work with the teacher is one of communication: interdisciplinary communication, and communication with administrators as well as with children.

On the question of the initial observation by the psychiatrists, and consultants, it was felt that it was not necessary for teams to visit together, but that it would be more efficient and easier to have a person from each discipline go at his own time and convenience, but quickly when called by the teacher. There was a felt need, however, for trained observations of the classroom process as a whole unrelated to specific individuals. This probably could be done in the beginning weeks of the school so that each class could be described as to its general tenor and group dynamics, and the techniques and style of the teacher. We also felt that we spent too much time on the individual and neglected seeing them in their total setting. We also felt that it was important to keep a variety of observers moving through the schools: psychologists, psychiatrists, teachers, etc. to help the teacher understand and get to know the different types of professions who work with children, and keep communication between disciplines open.

Our hypothesis that structure could be imposed on clinical and observational procedures to facilitate the inferential processes was strengthened with our experimentation in the three day classroom setting. By putting the children through specific tasks, making uniform demands on all the children, by testing specific personality functions, such as frustration tolerance and delay of gratification, we were able to begin standardizing, for ourselves at least, clinical and observational procedures. We set up a series of events, trained observers and discussed among ourselves the areas of personality we were specifically investigating. In general, the first day was usually permissive, letting the child explore the rooms, sharing in games, sitting together for snacks. The second and third day more demands were made of the children. Tasks were assigned and pressure put on the child to complete them. However, since the shortness of time was so pressing, and because we could get most of the reactions we wanted during the two-day sessions, we would recommend that the diagnostic classroom be limited in the future to two days, and then spend the entire third day in group discussion with teachers, social workers and the professional staff to work out the time-consuming final reports.

It was also felt that many more children were seen and tested in the school situation than actually came into the diagnostic classroom. These children should have a thorough work-up and recommendation as the ones who came into the diagnostic classroom. As it is now these children are not being followed-up because of the lack of time. One idea that has been discussed was to have a diagnostic classroom in the community so that each school with its own indigenous problems could have the diagnostic classes in situ. However, this would present the problem of handling observers. There were often 15-20 professional observers in the observational booth at the University. No school in the inner city is equipped to accommodate this number. We did feel that the numbers and the exchange they provided were beneficial in diagnosing the children and fulfilling the aims of the program. There might be a possibility of having a large trailer which would go from school to school; but, then again, the observation facilities would be limited.

We were confirmed in our hypothesis that in certain situations, it was more efficacious to have professionally trained workers rather than idigenous, nontrained aides. One of the major areas of difficulty was in getting information on the families. In the suburbs where trained social workers collected data, the information came through rapidly and competently. All the areas were covered that we wanted to know about, although there is always the question of the validity of parents reports.

However, we were looking not so much for specific information on the child as the attitude of the parents toward the child and his development, and this did come through in responses to our open ended questions. In the inner city the situation was different. We had hoped originally to elicit the aid of the neighborhood aides, that are local residents, so that we could help the community recognize the general problem of emotional disturbance in children and give them some training in interviewing. But the SNAP administration felt this to be too ticklish a situation to handle at this time. The neighborhood workers were unwilling to ask for this kind of information because it was too personal. They feel that they are interfering and prying. Also the information was open to misinterpretation by untrained workers. Consequently, on the advice of the senior social worker at SNAP, we let him do all the interviews. However, his schedules could not be maintained so he had to assign one of his neighborhood aides to collect the data. This information was scant and tantalizing. For instance, on a question on a child's relationship to the mother, the answer was "yes". Recorded trauma to the child, such as death of a parent or major injury, would be briefly recorded with no dates or any clue as to how this had effected the general family constellation or the development of the child. We still feelthat it would be a good idea to have the neighborhood aides learning and acting and supporting the program. We thought of having a trained social worker meet once a week with the neighborhood aides to help them learn how to collect data. SNAP discouraged taking forms in to be filled out by the families as this was too reminiscent of social welfare workers practices, which leads to resistance on the part of the parents. The aides, in general, know the families well and can give information without even asking specific questions of the family, as to the general tenor of the home, whether the mother's working, how many extended family are present, etc. This was the type of information we were looking for but which we are still trying to get. If we cannot get the aides to collect data and if we cannot train them ourselves, it would seem wise to have a professional social worker approach the problem of recording this information which is too important to have neglected or done haphazardly.

We also felt that the neighborhood aides could be used perhaps not to get all the information but at least to do a public relations job for us. So many of the parents had no idea of what was happening. Communications broke down very quickly between SNAP and parents. Many of the parents did not know that their children were coming to the University for diagnostic purposes and certainly did not know what we were doing at the University. One mother felt that we had given the children "needles" because they had been sick on their return from the University. This misconception could have been prevented if the neighborhood workers had been informing the parents about the program, making them see the usefulness of this program to their children. If we could have had a self-generating intercommunity program involving everyone in the community, we would have had greater success, not only in getting the diagnostic information but also in maintaining our follow-through study. We felt that the neighborhood aides should have been asked to the University. They should have been able to observe the sessions, and we should certainly have gone to their meetings and worked with them on their own specific problems. The main idea would be to get the community to work for itself, have the neighborhood workers give us gross information about the families, and be a public relations person for the program, and let the probing interviews and delicate subjects be approached by professionals not living in the immediate neighborhood.

The psychological testing raised several questions. Here it was felt that a professional, not a student in training, was needed to get the information from these young children. However, we would consider doctoral degree candidates with experience and training in this area as possible testers for the program. Although it was comparatively easy to test the children in the suburbs in one session, this was decidedly not so in the inner city. You needed two to three sessions with the child to get a valid sample of his performance. The child in the inner city initially resisted the meetings of the tester, but by the second or third day, he was quite willing to cooperate and his score was quite notably higher. The tester should also visit in the classroom for long observational periods to get to know the children and to look at cognitive function expressed in behavior in the classroom. All of this requires a sophisticated tester. If it is a student who is doing the testing, he should be closely supervised. There was some question whether the Binet was the appropriate test to use. It was suggested that perhaps a series of specialized tests and the standardizing of the observation and psychological approaches would be more effective in predicting emotional disturbance than the Binet alone. We would like to have tests of the relationship of the child with the teacher and with his peers and we would like to know more of his cognitive style and particularly any kind of tests, such as the Bender or equivilant that would reveal any central nervous system deficit or minimal brain disfunction.

We would like also to have noted changes in the classroom; that is, were there any differences in cognitive style as the child approached diverse tasks and added stress; what were the effects of different teaching styles and demands on different cognitive styles. The attitude of the child towards education and towards white people should have been noted to estimate of the child's level of aspiration and his ability to take risks. These are the data we felt would be helpful to describe children. Also included in this list of areas to explore were the child's handling of failure and his expression of aggression. Could the child request help from the peers and adults? Could he follow suggestions when they were given? Could he follow directions? Was there any relationship between social economic status on his ability to relate to peers and adults and follow directions? Could some of Piaget's and Brunner's tests be adopted to this population to see how cognitive hierachies were developing and how these hierachies were intergrated? Could special teaching techniques accelerate lagging areas of development? What types of materials interested the child? Did the child manifest a concrete or abstract approach to materials. Was he rigid or flexible in following the daily program? Was rigidity a form of perservation or obsessive behavior? How appealing was the child in general? How was the impulse control of the child? How as the child able to use play, for pleasure, mastery, learning of skills, people, things? Could he sustain what he started? What developmental tasks was he involved in? Could most of this material be gathered by trained observers-recorders watching free play?

On the subject of free play, we also wanted to see the sex role and identification processes, the choices of play material, whether they were male or female categories. In the classroom we would try to standardize the equipment and materials available, as well as the sequence of presentation of these materials so that we could have standardized means of analyzing the evoked behavior.

Again, this raises the issue of training. If the observers-recorders were undergraduates or people with little training in child development, cognitive process and clinical psychopathology, they are not able to observe with any kind of efficiency or report the significant material. Therefore, these observers-recorders should be first or second year master degree students or doctoral candidates who have been specifically trained as observer-recorders in a classroom for emotional disturbance.

BIBLIOGRAPHY

Bennet, William S., Jr. and Gist, Noel P., "Class and Family Influences on Student Aspirations," *Social Forces* 43 (1964) 167-173. (Aspirations and plans showed little variation according to class. Occupation aspirations show significant variance. Maternal influence stronger in lower classes regardless of race.)

Burgess, M. Elaine. "Some Implications of Social Change for Dependency among Lower-Class Families," *American J. of Orthopsychiatry* XXXIV (1964) 895-906. (Indicates Negro and white families on A.D.C. need more intensive aid than is now available.)

Charry, Mary A. W. *School Disorder, Intelligence and Social Class*. New York: Teachers College Press, 1967.

Dairs, Allison. *Social Class Influences upon Learning* (1948).

Deutch, Martin. "The Role of Social Class in Language Development and Cognition," *American J. of Orthopsychiatry* XXXV (1965) 78-88. (Low economic status is equated with low language development, becomes more marked in 5th rather than 1st grade.)

Dunham, H. "Social Class and Schizophrenia," *American J. of Orthopsychiatry* XXXIV (1964) 634.

Harrison, S. J., et. al. "Social Class and Mental Illness," Arch. Gen. Psychiat. (Nov., 1965).

Hernandez, David E. "Is the Concept of Social Class Being Misused in Education?" *Journal of Educational Sociology* 36 (1963) 322-324. (Concludes that high school teachers have better communication with lower-class parents.)

Hollingshead, A. B. *Elmtown's Youth*. New York: John Wiley and Sons, Inc., 1949.

Hollingshead, August B. and Redlich, Frederick C. *Social Class and Mental Illness*. (New York: John Wiley and Sons, Inc. 1958).

Jensen, Arthur R. "Social Class and Perceptual Learning," *Mental Hygiene*, 50 (1966) 226-239. (Argues school achievement differences by social class related to enrichment and impoverishment. Criticizes readiness.)

Luchterhand, Elmer and Weller, Leonard. "Social Class and the Desegregation Movement: A Study of Parents' Decisions in a Negro Ghetto," *Social Problems* 13 (1965) 83-88. (Found more school transfers in lower class and more involvement in organizations.)

McDermott, John F.; Harrison, Saul I.; Schraga, Jules and Wilson, Paul. "Social Class and Mental Illness in Children: Observations of Blue-Collar Families," *American J. of Orthopsychiatry* XXXV (1965) 500-508. (Found significant differences between children of skilled and unskilled parents in school performance and personality disorders.)

Miller, Daniel R. and Swanson, Guy E. In collaboration with Wesley Allinsmith and others. "Inner Conflict and Defense". (New York: Holt, 1960)

Miller, Kent S. and Grigg, Charles M. (eds.) Mental Health and the Lower Social Classes (Tallahassee: The Florida State University, 1966) (Collection of articles--none directly relevant to education of parental participation.)

Sears, Robert R.; Maccoby, Eleanor E. and Levin, Harry. Patterns of Child Rearing (White Plains, New York: Row, Peterson & Co., 1957) (Interview schedule and rating scales for mothers very helpful, located in Appendixes.)

Shoemaker, Louise Proehl. Parent and Family Life Education for Low-Income Families (U.S. Dept. of Health, Education and Welfare, 1965) (Discussion of problems in reaching low-income families--largely for social work point of view.

Swift, D. F. "Social Class and Achievement Motivation," Educational Research 8 (1966) 83-95. (Review of research.)

PERCEPTION AND SOCIAL
VALUES

HEAD START EVALUATION AND RESEARCH CENTER

Perception of Emotion Among Children: Race and Sex Differences

A. George Gitter

Boston University

ABSTRACT

This report describes a three phase study of children's perception of emotion investigating the patterns of accurately perceived and erroneously perceived emotions and testing the influence of sex and race differences among both perceivers and expressors.

---

*The research reported herein was performed pursuant to a contract for the Office of Economic Opportunity, Executive Office of the President, Washington, D.C. 20506. The opinions expressed herein are those of the author, and should not be construed as representing the opinions or policy of the United States Government.

HEAD START EVALUATION AND RESEARCH CENTER

PERCEPTION OF EMOTION AMONG CHILDREN: RACE AND SEX DIFFERENCES*

A. George Gitter

Boston University

Much of our daily social interaction involves nonverbal communication (NVC). Although NVC is a ubiquitous phenomenon, it has not previously received the attention paid to other areas of psychology. A number of questions arise pertaining to NVC which are of considerable interest toward understanding the mechanisms underlying social interactions, whether in the classrooms, in play groups, in occupational or other settings. Is race a significant variable in perception of emotion (POE)? If there are racial subculture differences in POE, are the cross-race and same-race POEs equal to each other? Are racial differences also significant in the expression of emotion? Are they significant in influencing the patterning of both accurately and erroneously perceived emotions? Are they equal in magnitude to sex differences in POE?

This study investigates the accuracy of perception of various emotions among children, as well as the patterning of erroneously perceived emotions. It focuses on the relationship of accuracy to race and sex differences of both the expressor (the person expressing the emotion) and the perceiver (the person judging the nature of the expressed emotion). There are a number of studies dealing with NVC in general, but only a few which relate to minority group characteristics or concentrate on race and sex variables in POE.

---

*The research reported herein was performed pursuant to a contract for the Office of Economic Opportunity, Executive Office of the President, Washington, D.C. 20506. The opinions expressed herein are those of the author, and should not be construed as representing the opinions or policy of any agency of the United States Government.

The most recent reviews of studies in perception, or recognition of emotion by means of NVC, are by Bruner and Tagiuri (1954), Davitz (1964), and Ekman (1965). These reviews indicate that posed photographs have been the most common stimuli used in the study of recognition of emotion by means of NVC cues. Other kinds of stimuli were utilized by various researchers, (Bruner and Tagiuri 1954), such as recording a person's voice, drawings of the human face with interchanging features, the use of people actually present, drawings of real persons, and photographs of a person accompanied by recordings of his voice.

A number of NVC studies arise out of psychiatric settings. Highly sensitive and affect-laden therapist-client interaction places heavy demands on the verbal channel, which in turn leads to the utilization of other (nonverbal) cues both for diagnosis and therapy. Mahl's (1956) investigation of the disturbances, discontinuities, and pauses of silence in the speech patterns of mental patients provides an example of such a study. He found significant correlations between these variables and the amount of anxiety present in the patient. Another study (Mahl, *et*. *al*. 1959), related gestures and bodily movements of psychiatric patients to their personality characteristics; its results indicated a significant relationship between tension and motor activity. The findings of Dittman (1962) validate those of Mahl's studies. He was able to relate hand movements to particular moods such as anger, gloom, and calm, and, although unsuccessful in relating anxiety to "linguistic" (verbal) behavior, he succeeded in discriminating between high and low conflict patients on the basis of their NVC (body movement) (Dittman and Wynne 1961).

Exline (1963), in exploring the function of visual cues, such as glances, during psychiatric interviews, found that stress inhibits eye-contact. In a replication study (Exline, *et*. *al*. 1965), his subjects, in an interview situation, evidenced substantially more eye-contact when they listened than whey they spoke. In addition, when the subject matter of the interview became embarrassing, the subject's visual contact was reduced to a minimum.

Judging minority group characteristics. Secord, Stritch, and Johnson (1960) have investigated the perception of personality characteristics through the use of nonverbal cues by offering their subjects limited stimulus information (photographs of faces). Their findings indicate that subjects tend to use some form of analogy, a "metaphorical generalization," or a bridge to fill in any information not supplied by the stimulus, and to organize all stimulus information into meaningful structures.

Secord, Bevan, and Katz (1956) and Secord (1959) investigated Negroes and Caucasians as stimuli. The subjects rated Negroes possessing pronounced Caucasian features with, "all the characteristics of a Caucasian," when restricted to the stimulus of a photograph and believing that the person in the photograph was Caucasian. On the other hand, other subjects on being informed that the photographs were of Negroes, rated these with "all the characteristics of a Negro."

Anisfield, Bongo, and Lambert's (1962) study investigated the relationship between perception of a minority group status and ascription of certain traits to such minority group members. The same actor taped two versions of a speech. In one version he used a "Jewish" accent, while in the other he did not. Those subjects who perceived him as "Jewish" rated him lower on such variables as: height, good looks, and self-confidence, while those who saw him as "English" rated him much higher on these same variables.

Race and sex differences. Gates (1923), working with white children perceivers and white adult female expressors, found age and sex of perceiver to be significant. Kellog and Eagleson's (1931) study of Negro children perceivers and white female adult expressors confirms the significance of sex differences. However Kellog finds no difference when he compares his results to those of Gates's white children. Studies have also found sex of expressor differences to be significant. (Thompson and Meltzer 1964; Drag and Shaw 1967).

Vinacke's (1949) study used magazine pictures of Caucasians as stimuli. His results indicated that Caucasian subjects, as compared to Oriental ones, were more accurate in identifying emotions of Caucasian faces. A replication study, by Vinacke and Fong (1955) similarly indicated that Oriental subjects, as compared to Caucasian ones, were more accurate in identifying emotions expressed in photographs of Oriental faces. In both studies, the magnitude of perceivers' sex differences was greater than that of race.

METHOD

Briefly, this study:

1. Investigates patterns of children's:
    - A. correctly perceived emotionns, and
    - B. erroneously perceived emotions (i.e., those which are in fact perceived, when they are not in fact expressed), and

2. Relates POE to the:
    A. race of (1) perceiver and (2) expressor, and
    B. sex of (1) perceiver and (2) expressor.

This research was divided into three phases.

Phase 1: Production of Stimuli materials--A period during which the feasibility of the various alternatives of data collection was explored and the stimuli materials were produced.

Phase 2: Efficacy of stimuli materials--A test of efficacy of the stimuli materials on a sample of adults before administering them to samples of children.

Phase 3: Investigation of POE among children--Design of research in progress, to be completed September, 1968.

Phase 1: Production of stimuli materials

July and August (1967), were spent investigating the feasibility of various data collection alternatives, in preparing stimuli materials, and making arrangements for samples of subjects, children of both races enrolled in Headstart Centers in the Boston area. My assistant, Mr. Harvey Black, has been extremely helpful with the former, while Mrs. Sandy Alexanian, of the School of Education, Boston University, was invaluable in making contacts for the latter.

Exploratory visits were made to Headstart Centers in three areas in the vicinity of Boston: Brockton, Revere and Roxbury. Tentative arrangements have been made to use the children in all of these centers as subjects in the study.

After considerable investigation of the relative advantages of the various types of stimuli materials, I decided to use motion pictures to photograph the enactments of the various emotions.

In part, the decision to use sound motion pictures was made because of its flexibility and potential use for future research, e.g., comparisons of POE by modalities (audio-only vs. visual-only vs. audio-visual). I was also influenced by the generous offer of Dr. G. Stechler of the B.U. Medical School to use his 16 mm. sound Auricon camera.

In order to obtain the necessary number of expressors, actors from local theater companies were contacted. A series of meetings to discuss the project were held with actors, mostly members of the Harvard Summer Players and the Peoples Theatre of Cambridge.

The motion pictures were made during a series of shooting sessions. A total of 398 takes were shot using 34 actors: 6 Negro males, 7 Negro females, 10 white males and 11 white females.

A rehearsal session was held preceding each shooting. At that time, the director, Mr. Vernon Blackman of the Peoples Theatre, discussed with the actors the emotions which they were to enact. Each actor acted out seven emotions, selected for ease of portrayal by the director: anger, mirth, surprise, fear, disgust, pain and sadness. Before each take, each actor again rehearsed the specific emotion to be filmed. The actors were filmed seated, facing 45° away from the camera, as if they were interacting with another person out of the view of the camera. Black and white film was used. Retakes were shot whenever either the director or the actor were dissatisfied for any reason with the original take.

As each actor enacted a particular emotion he recited a short monologue. The monologue was the same for all actors (two sentences) and all emotions, thus maintaining the verbal (or semantic) content constant across all of the experimental treatments.

To collect a series of "candid" still pictures of the seven emotions a search of Ebony, Life, and Look was undertaken. The paucity of results of this relatively extensive search led to the decision to utilize "posed" photographs, stills made from our motion pictures. A panel of three graduate student judges examined the total foorage for each emotion of each actor and actress. In each case, the frame judged representative of the emotion was used to make the still photographs.

Thirty-two hundred feet of motion picture film (approximately 90 minutes running time) has been processed. Mr. John Geeza of Magna Films, Inc., and Mr. Ray Richardson of Film Division, Boston University, edited the film footage.and made photographs of the frames selected from the motion picture films.

In the interests of balance of race and sex representation and consistency of photographic and dramatic quality, all of the motion picture material was examined. A final set comprising of emotions each for 12 males (6 white and 6 Negro) and 12 females (6 white and 6 Negro) was selected.

Phase 2: Efficacy of stimuli materials

To manipulate both the sex and race of the expressor a new set of stimuli materials was prepared. Before administering this set of photographs to children subject it was tested on a sample of adults in order to ascertain whether it compared in efficacy with the stimuli materials of other investigators. A copy of the text of the report of this study, submitted for presentation at the forthcoming (April, 1968) Eastern Psychological Association meetings, follows:

PERCEPTION OF EMOTION: RACE AND SEX DIFFERENCES OF EXPRESSORS AND PERCEIVERS

A 2x2x2x2 factorial study tested the effects of race and sex of perceivers (Ss, N=80), and race and sex of expressors (actors, N=24) each photographed enacting seven emotions. Race of expressor was found significant as were the patterns of both correctly and erroneously perceived emotions.

PROBLEM

Much of our daily social interaction involves the utilization of nonverbal communication in perception of emotion. Yet a number of questions, for example, (1) Those involving the effects of such variables as sex _and_ race of _both_ expressor and perceiver, or (2) the patterning of both correctly _and_ erroneously perceived emotions when the above variables are taken into account, have received scant attention.

SUBJECTS

Eighty (80) undergraduates made up the sample of perceivers (for the sex and race breakdown of the sample see the Experimental Design diagram below).

PROCEDURE

A balanced 2x2x2x2 design (below) tested the influence of the sex and race of both the expressor (the person portraying the emotion) and the perceiver (the person making the judgment as to the nature of emotion) on perception of emotion.

The twenty four (24) expressors in this experiment were professional actors (12 white and 12 Negro, 6 male and 6 female of each race--rows #1 and 2 of the Experimental Design diagram below).

1	Race	W								N							
2	Sex of Expressor	M				F				M				F			
3	Race of Perceiver	W		N		W		N		W		N		W		N	
4	Sex of Perceiver	M	F	M	F	M	F	M	F	M	F	M	F	M	F	M	F
5	N (for perceivers in each treatment group)	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5

w = white N = Negro m = male F = female

Black and white photographs were made of each expressor portraying seven (7) emotions: anger, happiness, surprise, fear, disgust, pain and sadness. Each photograph shows an expressor 3/4 figure, seated and at 45° angle to the camera. Lighting and background are constant for the 168 photographs.

The perceivers were the eighty (80) undergraduate Ss (40 white and 40 Negro, 20 male and 20 female of each race--rows #3 and 4 of the Experimental Design diagram). Each S was tested individually, with a set of twenty one (21) photographs (3 actors -- or actresses -- 7 emotions for each), presented in random order. After looking at each photograph the S made the judgement from a multiple choice list of the seven emotions (in alphabetical order).

RESULTS

1. Expressors and perceivers: race and sex differences

A. Neither sex nor race of the perceiver (Tables 1 and 2)* nor sex of the expressor (Table 3) influence the overall accuracy of perception of emotions.

B. Race of the expressor effects the accuracy of perception of emotion (Table 4), with higher accuracy occuring when perceiving white expressors as compared to their Negro counterparts. The extent of this influence however, is mediated by the sex and race differences of both the expressor and the perceiver. For example:

(1) While no differences were noted when overall accuracy scores were compared for perception of:

- (a) white male expressors versus Negro female expressors (Table 5), and
- (b) white female expressors versus Negro female expressors (Table 6),

(2) Clear differences emerge when accuracy comparisons involve:

- (a) Negro male versus Negro female expressors (Table 7)
- (b) white male versus Negro male expressors (Table 8), and
- (c) white female versus Negro male expressors (Table (9)

2. Patterns of correct perceptions

A. The incidence of correct perception varies with emotion (Table 10)--happiness and anger giving the highest, and fear and disgust the lowest proportions of correctly perceived emotions.

B. This pattern differs (Table 11) depending on whether the expressors are male or female.

C. No differences in the pattern of correct perception of the seven emotions appear as a function of the race of the expressor

---

*See the "Tables" page following Phase 2.

(Table 12) or the race and sex of the perceiver (Tables 13 and 14).

3. Patterns of erroneous perception. Erroneously perceived emotion is one, which was in fact perceived, whenever the perceiver made a mistake, that is, did not name the emotion enacted by the expressor in the photograph.

A. The incidence of erroneous perception varies with emotion (Table 15)--surprise and disgust giving the highest and pain and fear the lowest proportions of erroneously perceived emotions.

B. The pattern of correctly perceived emotions is different from that of erroneously perceived ones (Table 16).

C. The patterns of erroneously perceived emotions differ with sex of expressor (Table 17) and race of perceiver (Table 18).

D. No difference in patterns results from differences in either sex of perceiver (Table 19) or the race of expressor (Table 20).

IMPLICATIONS AND CONCLUSIONS

Sex and race differences among expressors and perceivers influence perception of emotion from posed photographs. While at times, the effects of these variables can be noted when their influence is examined independently, at other times they are significant only in interaction with each other. Similar results emerge from findings of the patterning of both correctly and erroneously perceived emotions.

Future research will hopefully answer whether the strong differences due to the race of expressor can be accounted for by white versus Negro subcultural differences in inhibiting the communication of emotions, or by such artifacts as the inability of black and white photographs to capture portrayal of emotion equally well for white and Negro expressors, or by still other alternative explanations.

TABLE 1*

	C	I	T
WP	474	366	840
NP	485	355	840
T	759	721	1680

$\chi^2 = .12$ df=1 N.S.**

TABLE 2

	C	I	T
MP	455	385	840
FP	504	336	840
T	959	721	1680

$\chi^2 = 2.74$ df=1 N.S.

TABLE 3

	C	I	T
ME	456	384	840
FE	503	337	840
T	959	721	1680

$\chi^2 = 2.30$ df=1 N.S.

TABLE 4

	C	I	T
WE	526	314	840
NE	433	407	840
T	959	721	1680

$\chi^2 = 9.03$ df=1 $p < .01$

TABLE 5

	C	I	T
WME	262	158	420
NFE	239	181	420
T	501	339	840

$\chi^2 = 2.63$ df=1 N.S.

TABLE 6

	C	I	T
WFE	264	156	420
NFE	239	181	420
T	503	337	840

$\chi^2 = 3.44$ df= 1 N.S.

TABLE 7

	C	I	T
NME	194	226	420
NFE	239	181	420
T	433	407	840

$\chi^2 = 8.08$ df=1 $p < .01$

TABLE 8

	C	I	T
WME	262	158	420
NME	194	226	420
T	456	384	840

$\chi^2 = 22.18$ df=1 $p < .001$

TABLE 9

	C	I	T
WFE	264	156	420
NME	194	226	420
T	458	384	840

$\chi^2 = 23.50$ df=1 $p < .001$

*CODE

C=CORRECT I= INCORREC
M= MALE F= FEMALE
W= WHITE N= NEGRO
E= EXPRESSOR P= PERCEIV

TABLE 10 EMOTIONS***

	1	2	3	4	5	6	7	T
All P-All E	160	212	128	77	114	149	119	959

$\chi^2 = 53.22$
df=
$p < .001$

TABLE 11

	1	2	3	4	5	6	7	T
ME	75	113	47	18	69	77	57	456
FE	85	99	81	59	45	72	62	503
T	160	212	128	77	114	149	119	959

$\chi^2 = 42.77$
df = 6
$p < .001$

TABLE 12

	1	2	3	4	5	6	7	T
WE	87	114	72	44	64	73	72	526
NE	73	98	56	33	50	76	47	433
T	160	212	128	77	114	149	119	959

$\chi^2 = 6.10$
df=6
N.S.

TABLE 13

	1	2	3	4	5	6	7	T
WP	77	109	66	36	52	72	62	474
NP	83	103	62	41	62	77	57	485
T	160	212	128	77	114	149	119	959

$\chi^2 = 1.99$
df= 6
N.S.

TABLE 14

	1	2	3	4	5	6	7	T
MP	78	104	62	31	51	74	55	455
FP	82	108	66	46	63	75	64	504
T	160	212	128	77	114	149	119	959

$\chi^2 = 3.23$
df=6
N.S.

TABLE 15

	1	2	3	4	5	6	7	T
Err All P-All E	128	86	167	53	137	52	98	721

$\chi^2 = 109.36$
df= 6
$p < .001$

TABLE 16

	1	2	3	4	5	6	7	T
All C	160	212	128	77	114	149	119	959
All Err	128	86	167	53	137	52	98	721
T	288	298	295	130	251	201	217	1680

$\chi^2 = 85.97$
df=6
$p < .001$

TABLE 17

	1	2	3	4	5	6	7	T
ME	80	48	82	17	92	26	39	384
FE	48	38	85	36	45	26	59	337
T	128	86	167	53	157	52	98	721

$\chi^2 = 29.40$
df=6
$p < .001$

TABLE 18

	1	2	3	4	5	6	7	T
WP	62	39	96	35	61	21	52	366
NP	66	47	71	18	76	31	46	355
T	128	86	167	53	137	52	98	721

$\chi^2 = 14.18$
df= 6
$p < .05$

TABLE 19

	1	2	3	4	5	6	7	T
MP	77	40	94	22	74	33	45	385
FP	51	46	73	31	63	19	53	336
T	128	86	167	53	137	52	98	721

$\chi^2 = 12.34$
df= 6
N.S.

TABLE 20

	1	2	3	4	5	6	7	T
W.E	62	33	67	29	60	24	39	314
NE	66	53	100	24	77	28	59	407
T	128	86	167	53	137	52	98	721

$\chi^2 = 6.77$
df=6
N.S.

** *All two tail tests*

***CODE

(1) ANGER
(2) HAPPINESS
(3) SURPRISE
(4) FEAR
(5) DISGUST
(6) PAIN
(7) SADNESS
Err= ERRONEOUS

Phase 3: Investigation of POE among children--design of research in progress to be completed September, 1968

Research Design

Race of expressor	W				N				
Race of perceiver(s)	W		N		W		N		
Sex of subject	M	F	M	F	M	F	M	F	
N = (in each group)	15	15	15	15	15	15	15	15	

w = white
N = Negro

M = male
F = female

Subjects. Children (age 5-6) of both races and sexes (N=120; 60 white and 60 Negro, 30 males and 30 females of each race), enrolled in Headstart programs, will be used as subjects.

Procedure. Still photographs of females (expressors) of both races will be used as stimuli materials. Each S will be asked to identify the emotion in each of 12 photographs--3 actresses, either white or Negro, depending on the treatment group, each photographed portraying 4 emotions (happiness, pain, anger and surprise). Ss will be tested individually and the order of presentation of photographs will be random for each S.

Upon being presented with a photograph the S will be asked which of the four emotions is depicted by the expressor. Because considerable amount of evidence from pretesting indicates that a five-year old's span of attention may limit his ability to keep the four alternatives (the four emotions) in mind, a training period will precede the testing session. Four comic book art type pictures will be placed on the table in front of the S, each depicting a scene typical of the four emotions. The E will explain to the S the four pictures--how each is to remind the S of each of the emotions. The pictures will remain on the table through the testing session allowing the E to refer to them. In short, they will act as graphic multiple choice alternatives.

REFERENCES

1. M. Anisfield, N. Bongo, and W. Lambert, "Evaluational Reactions to Accented English Speech," *Journal of Abnormal and Social Psychology*, *65*, 1962, pp.223-231.

2. J.S. Brunner and R. Tagiuri, "Tagiuri, "The Perception of People," in G. Lindzey (Ed.), *Handbook of Social Psychology* (Cambridge, Mass.: Addison-Wesley, 1954), pp. 635-654.

3. J.R. Davitz, *The Communication of Emotional* Meaning (New York: McGraw-Hill, 1964).

4. A.T. Dittman, "The Relationship Between Body Movements and Mood in Interviews," National Institute of Mental Health, 1962, pp. 1-10.

5. A.T. Dittman and L.C. Wynne, "Linguistic Techniques and the Analysis of Emotionality in Interviews," *Journal of Abnormal and Social Psychology*, *63*, 1961, pp. 210-204.

6. R.M. Drag and M.E. Shaw, "Factors influencing the communication of emotional intent by facial expressions," *Psychon. Sci*., 1967, *18*(4) pp. 137-138.

7. P. Ekman, "Communication through Nonverbal Behavior: A Source of Information about an Interpersonal Relationship," in S. Tomkins and C. Izard (Eds.), *Affect, Cognition, and Personality* (New York: Springer, 1965), pp. 390-443.

8. R.V. Exline, "Explorations in the Process of Person Perception: Visual Interaction in Relation to Competition, Sex, and Need for Affiliation," *Journal of Personality*, *31*, 1963, pp. 1[illegible]-20[illegible].

9. R.V. Exline, D. Gray, and D. Schuette, "Visual Behavior in a Dyad as Affected by Interview Content and Sex of Respondent," *Journal of Personality and Social Psychology*, *1*, 1965, pp. 201-209.

10. G.S. Gates, "An experimental study of the growth of social perception," *Journal of Educational Psychology*, 1923, *14*, pp. 449-461.

11. W.N. Kellogg and B.M. Eagleson, "Growth of social perception in different racial groups," *Journal of Educational Psychology*, 1931, *22*, pp. 367-375.

13. G.F. Mahl, B. Danet, and N. Norton, "Reflection of Major Personality Characteristics in Gestures and Body Movements," Paper read at the American Psychological Association, Annual Meeting, 1959, pp. 1-18.

14. P.F. Secord, "Stereotyping and Favorableness in the Perception of Negro Faces," Journal of Abnormal and Social Psychology, 59, 1959, pp. 309-314.

15. P.F. Secord, W. Bevan, and B. Katz, "The Negro Stereotype and Perceptual Accentuation," Journal of Abnormal and Social Psychology, 53, 1956, pp. 78-83.

16. P.F. Secord, T.F. Stritch, and L. Johnson, "The Role of Metaphorical Generalization and Congruency in the Perception of Facial Characteristics," Journal of Social Psychology, 52, 1960, pp. 329-337.

17. D.F. Thompson and L. Meltzer, "Communication of emotional intent by facial expression," Journal of Abnormal Social Psychology, 1964, 68, No. 2, pp. 129-135.

18. W.E. Vinacke, "The Judgment of Facial Expressions by Three National-Racial Groups in Hawaii: I. Caucasian Faces," Journal of Personality, 17, 1949, pp. 407-429.

19. W.E. Vinacke and R.W. Fong, "The Judgment of Facial Expressions by Three National-Racial Groups in Hawaii: II. Oriental Faces," Journal of Social Psychology, 41, 1955, pp. 185-195.

HEAD START RESEARCH AND EVALUATION CENTER

THE EXPRESSION OF AGGRESSION IN PRE-SCHOOL CHILDREN[1]

LYNN M. DORMAN

BOSTON UNIVERSITY

ABSTRACT

One behavioral and two projective measures of aggression were given to preschool children enrolled in two Head Start summer sessions in Quincy and Lynn, Massachusetts. The projective measures were designed to elicit the expression of aggression in social situations. The situations were swinging, painting, playing with blocks or dolls and holding a book.

The frequency of aggressive responses given by the children was computed. There was no difference in the amount of expression of aggression in the two schools, but the girls in Quincy gave more aggressive responses than any of the other children. Possible factors relating to this were discussed. These included verbalness of females, greater familiarity with the experimenters, racial and age differences.

The behavioral measure did not elicit any overt aggression and the responses were coded for verbal and cooperative responses between the two partners. The responses of those children ranking high and low in the expression of aggression were compared. Children who expressed the most aggression were more verbal and more cooperative than those who expressed little or no aggression. This finding was discussed in terms of the relation between learning ability and the controlled expression of aggression.

The relationship between expression of aggression and ability to learn is the subject of a study currently being planned by the author. If this relationship can be shown, it will have implications for the kinds of programs set up for those children who have learning difficulties. The freeing of appropriate aggressive responses might lead to increased freedom to deal appropriately with other aspects of the environment, such as the learning situation.

---

[1] "The research reported herein was performed pursuant to a contract with the Office of Economic Opportunity, Executive Office of the President, Washington, D.C., 20506. The opinions expressed herein are those of the author and should not be construed as representing the opinions or policy of any agency of the United States Government."

THE EXPRESSION OF AGGRESSION IN PRE-SCHOOL CHILDREN[1]

LYNN M. DORMAN

BOSTON UNIVERSITY

In the pre-school years, the child's interaction with peers increases and his mode of interaction changes. As he proceeds from solitary to cooperative play (Parten, 1943), he develops patterns of responding to and in these social situations. These response patterns include the expression of cognitions and emotions appropriate to the situation. The development of these response patterns at this time form the basis for other more sophisticated patterns of responses at a later stage of development. The expression of aggression is only one of these response patterns, but is an important factor in the child's socialization and in his general development.

Aggression is a natural human emotion which needs appropriate outlets (Bettelheim, 1966). Aggression can be dichotomized into 1) instrumental aggression: actions such as manipulating the environment and asking questions, 2) non-instrumental aggression: actions such as striking out at someone and damaging property (Beller, 1957). The management of aggression through controlled appropriate expression may be viewed as one of the precursors to the more complex response of assertion which is necessary for social development in general and for learning in particular.

The purpose of this study was to develop an instrument for measuring the expression of aggression by pre-school children in social situations. Pictures, dolls and hypothetical situations were initially used as projective measures of aggression. A measure of overt aggression using a block game was added later in the study.[2] It was hoped that such instruments could be later used in a study of the expression of aggression as related to learning in pre-school children.

The present study was exploratory. No working hypotheses were developed prior to the study, but the findings suggested a hypothesis for further study.

METHOD

Subjects: The subjects were children from two classes in each of two Head Start populations. The schools were in Quincy and Lynn, Massachusetts. The children in the Quincy school were all white and of pre-kindergarten age. The atmosphere of the two Quincy classrooms was one of general permissiveness during free play at which time observations were made. There were 9 boys and 4 girls in one room and 8 boys and 6 girls in the other. The Lynn, Massachusetts, population consisted of both negro and white children who were to enroll in the first grade in the fall. The children were children were more restricted, working at their desks, and were often engaged in specific learning tasks. These children were taken on many school sponsored trips during the summer. The make-up of one room was 6 boys and 4 girls; the other had 6 boys and 9 girls. The general family background (income and number of children) was the same for both populations. Due to absences in Lynn and refusals in Quincy, complete data was collected on 6 boys and 8 girls in Quincy and 4 boys and 8 girls in Lynn. The children who refused to look at the pictures or dolls seemed to be those who were having trouble adjusting to school and were "quiet," "shy," and cried during the first few days of school.

---

[1] "The research reported herein was performed pursuant to a contract with the Office of Economic Opportunity, Executive Office of the President, Washington, D.C., 20506. The opinions expressed herein are those of the author and should not be construed as representing the opinions or policy of any agency of the United States Government."

[2] A complete description of the materials used will be found in the appendix.

Procedure: Due to the classroom differences, testing time in Quincy extended over a longer period of time than in Lynn. The experimenters were in Quincy the first day the class met so that they were more familiar to these children than they were to the Lynn children. The children were seen individually within the classroom, and ongoing activities and play were not interrupted. Each child was asked to respond to three social situations depicted by dolls and drawings. The hypothetical question involved the child placing himself into the three situations. The three social situations involved block play and swinging for boys, doll play and painting for girls, and a book situation for both sexes. These situations were chosen because they reflected popular sex appropriate activities for this age. Each situation portrayed two same-sexed children, except for the book which had one boy and one girl. The situations could elicit an aggressive or a non-aggressive response. An aggressive response was defined as one in which one of the figures did something to the other or to the other's possessions such as pushing off the swing, hitting, or knocking over the blocks. Any other response was considered a non-aggressive response. The child was first asked if he wanted to see some pictures or dolls. After the child sat down, the experimenter brought out the pictures or dolls and said to the child, "Tell me what is happening." An example of the hypothetical questions is "what if you were playing dolls and another girl came over, what do you do?"

One experimenter worked with the pictures and another with the dolls. Children were never asked to respond to more than three situations at a time and were not presented with the dolls and pictures on the same day.

After all, the children had been questionned about the social situations, they were asked to play a block game in pairs. Any two children who volunteered to come at the same time were allowed to play; no attempt was made to systematically pair the children. The instructions were: "If you two can put all the blocks on one building, you each win a balloon. You have to take turns though. (First pointing to one child and then to the other) You put one on; then you put one on. You get three chances, but if all three buildings fall down, no one gets a balloon." The experimenter watched and noted what was said and done in the game. (The game used was Blockhead.[3]) It is virtually impossible to put all the blocks on one building and since no child won a balloon, the experimenter gave all children in the classroom a balloon.

As another measure of overt aggression all teachers were asked to rank order the children in terms of their display of aggression in the classroom. The children were rated during the last week of the summer session.

---

[3] See appendix

Results:

All responses were coded for aggression or non-aggression as defined above. Examples of aggressive responses are: "he pushed him off the swing," "they are fighting over the book;" non-aggressive responses were: "they are playing dolls nicely," "the paint spilled accidentally." No one expressed aggression during the block game; and therefore, it was coded for verbal or non-verbal interaction and for cooperation or non-cooperation during the three turns. Cooperation was considered any attempt to influence the other player, i.e., suggesting how to build, planning ahead, pointing at a block the other should use.

The results to be reported here are preliminary. More extensive data analyses are in progress. Long range analyses will compare the child's responses from task to task and within situations to determine whether the type of task or the situation calls forth differences in the expression of aggression. The following analysis is based on the responses to the pictures and the dolls. The analyses were of six responses (3 responses to pictures and 3 responses to dolls) instead of the nine elicited. The hypothetical questions were not included because it was difficult to standardize the questioning. All the children were asked the original questions, but if they gave no answer, the experimenter would reword the question suggesting that the hypothetical person in the situation came over to aggress. There is a possibility that this may be a useful measure as it has the potential for eliciting "true-to-life" responses.

Table I shows the frequency distribution of aggressive responses for all children. The greatest number of responses each child could give was six (3 for the picture and 3 for the dolls).

TABLE I

DISTRIBUTION OF AGGRESSIVE RESPONSES

Number of Aggressive Responses	Quincy		Lynn	
	Male	Female	Male	Female
0	1	0	0	1
1	1	0	0	0
2	2	2	1	1
3	2	1	0	3
4	1	2	0	2
5	0	0	2	0
6	0	3	1	1

All children except two expressed aggression. The task does allow for the expression of aggression. A greater sex difference was found in the Quincy population; the girls expressing the most aggression. In order to ascertain if there are differences between those who scored high on expressed aggression and those who scored low, the block game responses were analyzed for those children who expressed less than two and more than five aggressive responses, 5 and 7 subjects respctively. For those who scored high on aggressive responses, 7 of the 15 possible responses in the block game were verbal and 8 of the 15 were cooperative; whereas for those who scored low on aggressive responses, 4 of the 21 possible responses were verbal and 6 of the 21 responses were cooperative.

The correlations between our ratings of aggression and the rankings from the teachers 'was approximately .20 indicating little agreement between the teacher's ratings of overt aggression and our measure of expressed aggression.

DISCUSSION

Although one might have expected the differences in the environment to produce differences in the total amount of expressions of aggression, none were found.

In Quincy, the girls expressed more aggression than did the boys--a finding not in aggreement with the literature. (Sears, 1951). Perhaps in a more permissive atmosphere, girls are able to express their aggression verbally and the boys express it motorically. In the more "formal" setting, the sex difference in expression is not as great. There are many other possible influences on this factor: age and racial differences, degree of familiarity with experimenters, and classroom atmosphere.

A comparison of the highest and lowest expressers of aggression indicates that children who express more aggression on projective tests are more verbal and more cooperative, at least in the game situation used here. It is possible that the highly verbal, cooperative child expresses more aggression because he is verbal, but this did not seem supported here, since low aggressive children were also verbal when telling stories.

It is felt that verbalization and cooperation are two important aspects of both social development and school learning. The child who can express aggression in controlled situations where it is likely to be appropriate is probably more able to socialize and learn. These children can relate cognitively to their environment. Those who could not see or express aggression could not verbalize or cooperate and may possibly have difficulty in social or learning situations.

CONCLUSION

The projective measures developed in this study appear to be indicative of the child's ability to express aggression, although they were not related to the teacher's rating of overt aggression. Perhaps the projective expression of aggression is related to instrumental aggression, as discussed earlier. The children who were not able to express aggression in the projective situations were those who were less verbal and less cooperative on a task that involved some foresight.

REFERENCES

Bellar, E. Kuno Dependency and the frustration-aggression hypothesis II. Paper presented at the 32nd meeting of the Eastern Psychological Association, 1961.

Bettelheim, B. Violence: A neglected mode of behavior. Annals of the American Academy of Political and Social Science, 1966, 364, 50-59

Parten, M. Social participation among pre-school children. Journal of Abnormal Social Psychology., 1933, 27, 243-269.

Sears, P. Doll-play aggression in normal young children: Influence of sex, sibling status, father's absence. Psychological Monograph, 1951, No. 6 (Whole No. 323).

APPENDIX

The following is a description of the pictures used in this study:

1. A boy is kneeling by some blocks and a building is ready to topple. Another boy is standing nearby and looking at the situation.

2. A boy is lying on the ground in front of a swing while another boy is standing slightly to the rear of the swing.

3. A girl is playing with dolls. The play area is "messy" and an arm is broken off one doll. Another girl is standing nearby.

4. A girl is at an easel with a paint brush in her hand. One can of paint is tipped and the paint is spilling. Another girl is standing near the easel with

5. A boy and a girl are standing and both are holding the edges of the same book.

The dolls used in the study are manufactured by Creative Playthings in Princeton, New Jersey. The dolls were put into the same situations as described above using a shoe box as a background.

The hypothetical questions dealt with these same situations and were generally worded: "What if you were (painting, playing with blocks) and another boy (girl) came over, what do you do?"

The block game is a commercial game called Blockhead, a Saalfield Artcraft product. The game consists of blocks of various shapes and colors.

TEACHING PROGRAMS AND
LEARNING

HEAD START EVALUATION AND RESEARCH CENTER

LANGUAGE PROJECT: THE EFFECTS OF A TEACHER DEVELOPED, PRE-SCHOOL LANGUAGE TRAINING PROGRAM ON FIRST GRADE READING ACHIEVEMENT[1]

Sandra Alexanian

Boston University

ABSTRACT

The purpose of the project was to initiate a Teacher-Developed Pre-School Curriculum to Facilitate Grade One Reading Success.

The gap between what many Headstart programs offer and the expectations of public schools receiving the alumni is often great. It was the hypothesis of this project that a community oriented readiness curriculum could be developed utilizing: 1) Community vocabulary and resource, 2) Knowledge of public school materials, vocabulary and expectation, 3) Skills which have a demonstrable effect on early reading success (National Grade One Reading Study).

Testing of experimental and central groups was done at the begining and end of a summer Headstart term. The instrument used was the Murphy-Durrell Reading Readiness Analysis.

The vehicles used to stimulate the teacher developed curriculum were: 1) Workshops for experimental teachers and their aides with specialists in curriculum areas (pre-school language, reading, childrens books, drama and community). 2) Classroom support in the form of materials and demonstration.

The project will be complete when the subjects are tested again in January 1968 in relation to their grade one reading achievement.

---

[1] "The research reported herein was performed pursuant to a contract with the Office of Economic Opportunity, Executive Office of the President, Washington, D.C., 20506. The opinions expressed herein are those of the author and should not be construed as representing the opinions or policy of any agency of the United States Government."

HEAD START EVALUATION AND RESEARCH CENTER

THE EFFECTS OF A TEACHER DEVELOPED, PRE-SCHOOL LANGUAGE TRAINING PROGRAM ON FIRST GRADE READING ACHIEVEMENT[1]

Sandra Alexanian

Consultant: Alice Crossly

Boston University

This is a summary of the project to date. Final testing will take place in January of 1968.

The purpose of this study was to determine the effects of a teacher developed language readiness curriculum on the grade one reading achievement of selected groups of children enrolled in a Headstart program.

It was hypothesized that children who receive instruction in various language skills will demonstrate significant gains in Grade one reading achievement when compared to children who do not receive language readiness training.[2] The experimental group consisted of three Headstart classes(N=35), while the control group consisted of two Headstart classes(N=25). All subjects were eligible to enter the first grade in September, 1967. Males and females were equally represented in both experimental and control groups. All classes were part of one funding agency; geographically the subjects were situated in four separate school districts.

The language project, and its rationale was presented to the entire teaching community at a regular in-service meeting. Three teachers volunteered to participate as experimental classroom teachers. The two control classrooms were chosen by the agency program director.

All test subjects were administered the Murphy-Durrell Reading Readiness Analysis[3] by qualified elementary reading personnel. Testing took place the first week of July and the third week of August. The test instrument has five subtests.

---

[1] "The research reported herein was performed pursuant to a contract with the Office of Economic Opportunity, Executive Office of the President, Washington, D.C. 20506. The opinions expressed herein are those of the author and should not be construed as representing the opinions or policy of any agency of the United States Government."

[2] National Grade One Reading Study(HEW),@1967.

[3] Murphy-Durrell Reading Readiness Analysis, Harcourt, Brace and World, Inc., N.Y., N.Y., @ 1965.

The four sub-tests outlined in the National Grade One Reading Study as predicting grade one reading facility were used.[4] Pre-testing was completed in two days, while post testing took one day to complete.

Two major approaches were utilized with experimental teachers and their aids. Nine, two hour, workshop sessions were held. Teachers and aides were paid, as consultants, to attend each session. With the assistance of six consultants[5] the teachers were encouraged to explore the problems of presenting academic material to pre-schoolers, as well as to examine community experiences and grade one, public school expectations. The differences between preschool experience and public school expectation presented the teachers with the workshop task. The workshop content is described in the agenda (see appendix).

The remainder of the intervention took the form of classroom visits to introduce new materials, both published and unpublished; demonstrations of workshop developed material; a model field trip; and a demonstration of dramatic play techniques.

In illustrating the results (see appendix) the experimental teachers are numbered 1,2 and 3, the control teachers 4 and 5. The implementation of the evolving curriculum was reflected in the individualistic teaching style of each teacher. Teacher #1 was well trained, professionally sophisticated and self assured. Her class was controlled in a good humored manner, and the children were always aware of teacher expectation. She effectively implemented workshop recommendations. Any number of complicated interventions might have been introduced in this setting. Teacher #2 was continually "explaining" classroom behavior in terms of her personal assumptions regarding the effects of home and community influences. Although she was well trained, and maintained a permissive classroom atmosphere, it was difficult to introduce small group activities. Audio-visual materials[6] proved to be very effective with these children and the teacher. Teacher #3 was overly concerned with the mechanics of teaching. The "appearance" of teaching materials was more important than their actual value to the children. These concerns kept the teacher busy making and doing things for the children. The behavior of teachers as a group was characterized by energetic involvement in the project.

The control teachers were also cooperative toward the project. Teacher #4 was seen briefly at the introduction of the testing period and did not remain to observe the test. However, teacher #5, whose population was semi-rural, expressed strong motivation to participate in the workshop, and expressed disappointment when she was not permitted to be an experimental group teacher.

---

[4] 1) Phoneme Test I, initial phoneme sounds
2) Phoneme Test II, initial and final phonemes and diagraphs
3) Capital Letter names
4) Lower Case letter names

[5] Dr. Alice Crossly, Mr. Albert Cullum, Dr. Helen Murphy, Mrs. Sandra Alexanian, and Mrs. Mary Brassard all of Boston University, School of Education. Mr. Kiyo Morimoto, Harvard University, Bureau of Study Council. Center assistance to project was Wilma Snowdon.

[6] Crossly, Alice, The Evaluation of the Effects of Lantern Slides on Audio-visual Demonstration of Word Elements, 1948.

When implementing a teacher-developed curriculum, you must attempt to account for variations in teacher competance when assessing the effects. This has been expressed by many of the sub-studies in the National Grade One Reading Study. The teacher must believe in and value the content before an intervention can become effective. The method of asking teachers to modify a curriculum to incorporate community characteristics, strengths and deficits leaves the impact and implementation on the shoulders of the teacher. This is in contrast to the introduction of pre-determined curriculum and materials.

Until the final testing[7] takes place in January, when the Headstart alumni is in Grade One, only an informal descriptive analysis of data can be made. At that time a more formal analysis of the effects of a "Teacher-Developed Pre-School Curriculum" can be evaluated. However, upon examining the results of pre and post testing it would appear that the experimental classes have made greater gains than the control classes. If these gains are stable and significant, very likely the experimental classes can look forward to a greater facility in early grade one reading.

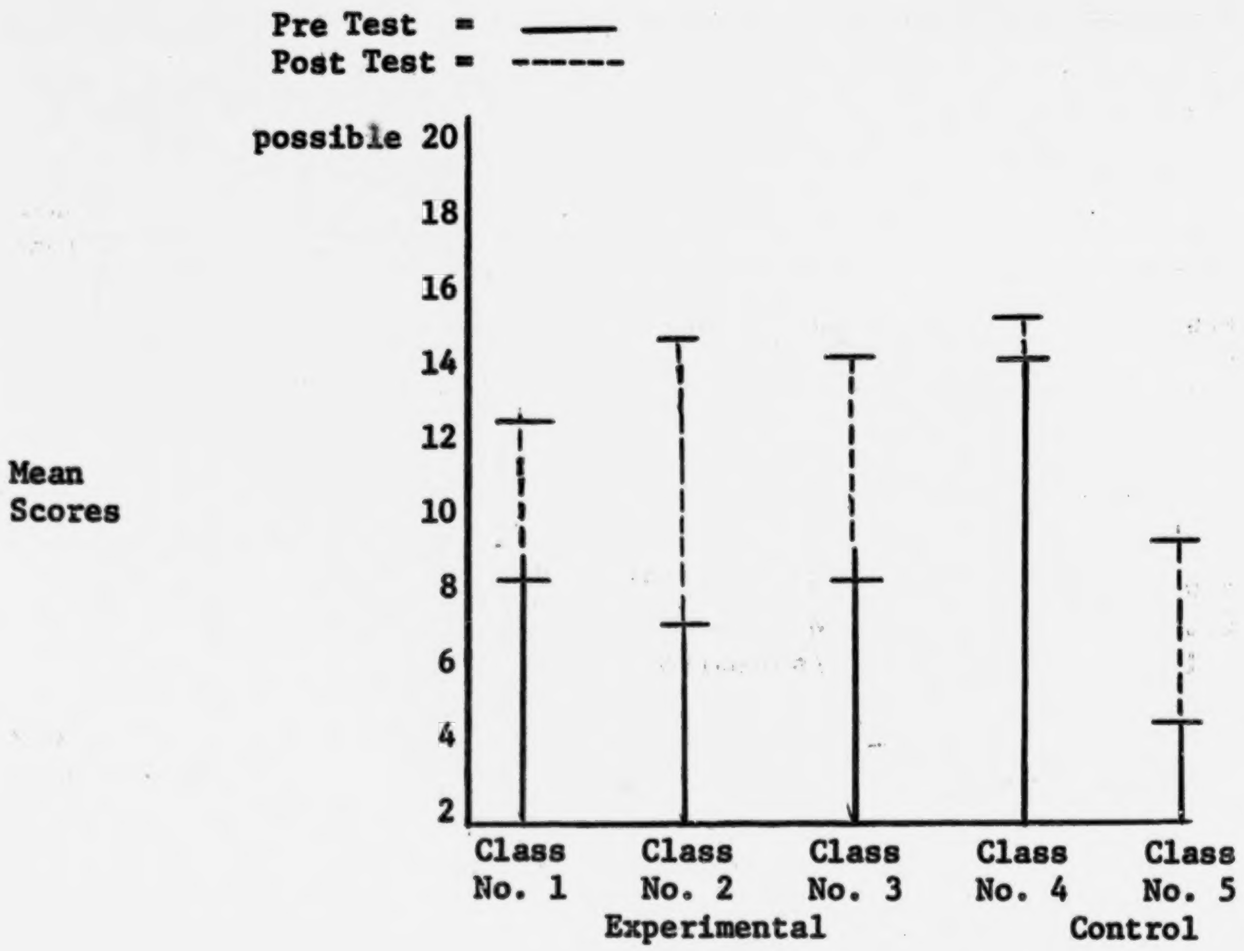
---

[7] Metropolitan Grade One Reading Achievement, Primary I, Harcourt, Brace & World, Inc.

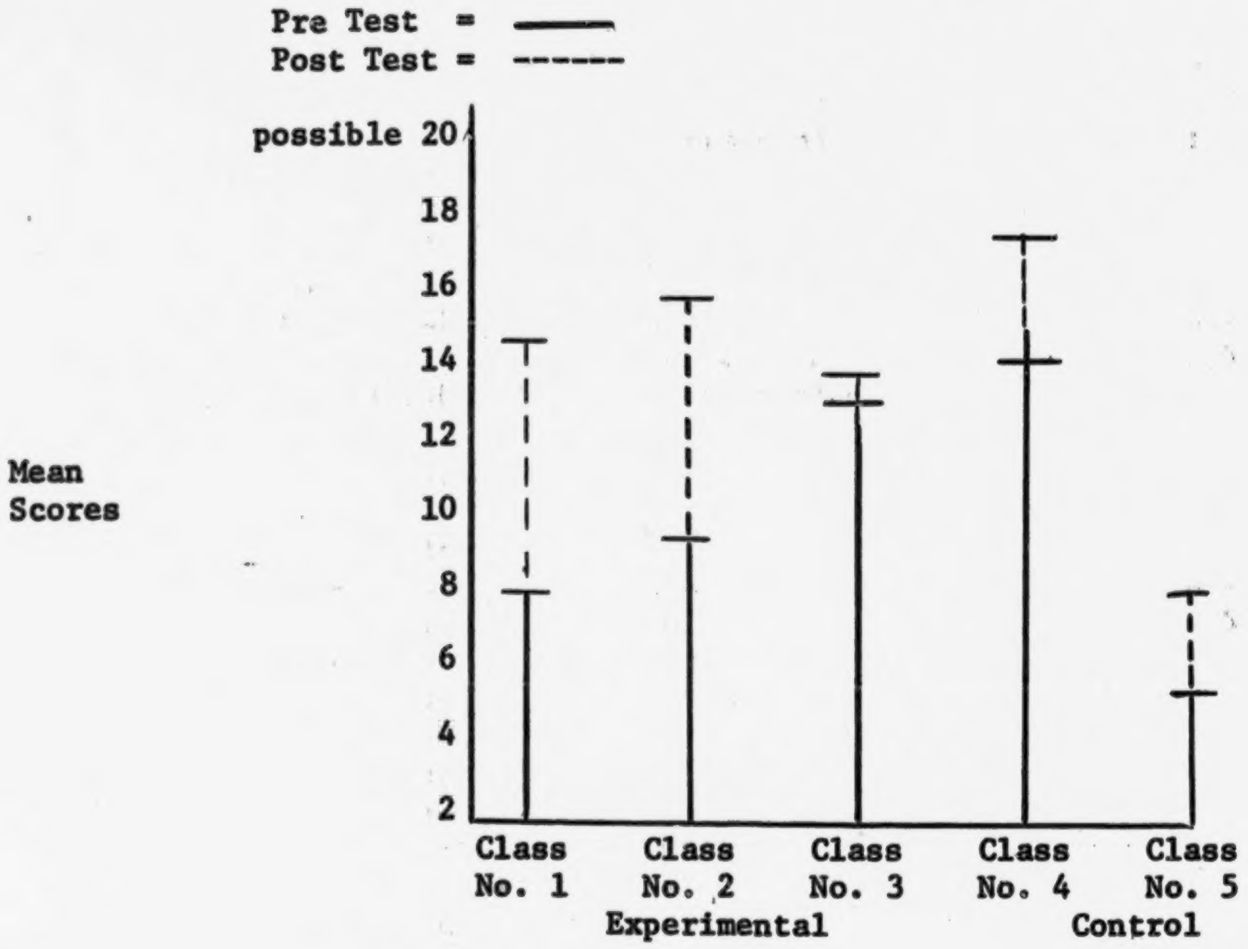
LANGUAGE WORKSHOP FOR HEADSTART TEACHERS

Date	Topic	Staff	Place
	Overview and Orientation	S. Alexanian Dr. Crossley	
	Grade One Expectations 1. Analysis of books for use 2. Analysis of concepts presented	Dr. Crossley	
	Community Resources 1. Real experiences possible 2. Community vocabulary	Dr. Crossley	
	Games and Devices for letter names and sounds	Dr. Helen Murphy	
	Report on the National Grade One Reading Study	S. Alexanian	
	Creative Drama	Dr. Crossley A. Cullum	
	Poems, Stories and Flannel Graph -vocabulary experiences	M. Brassard	
	Vocabulary Lessons Classroom Organization	Dr. Crossley	
	Grade 1 Instructional Vocabulary Working in Community	S. Alexanian Kiyo Morimoto	

PHONEMES I

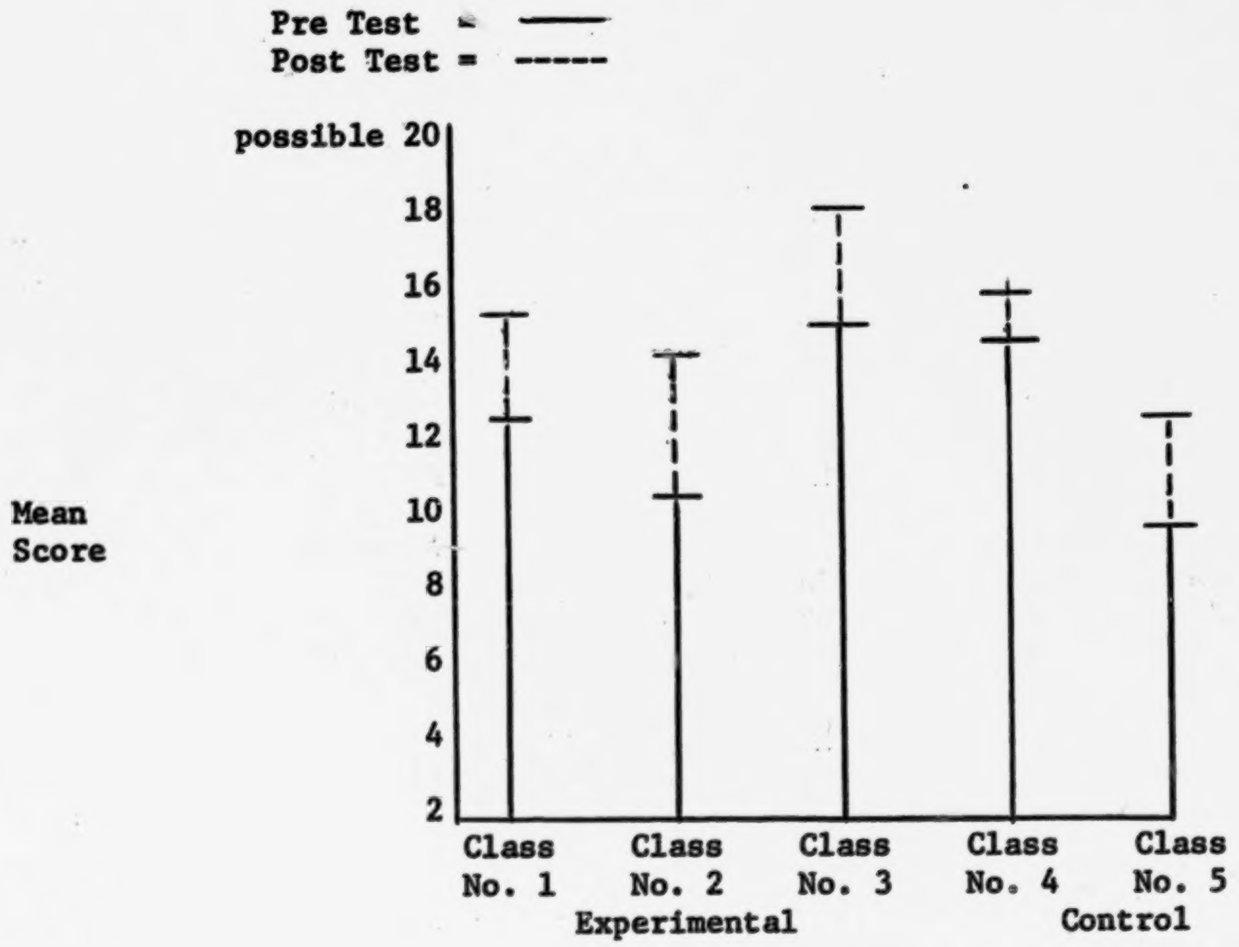


PHONEMES II

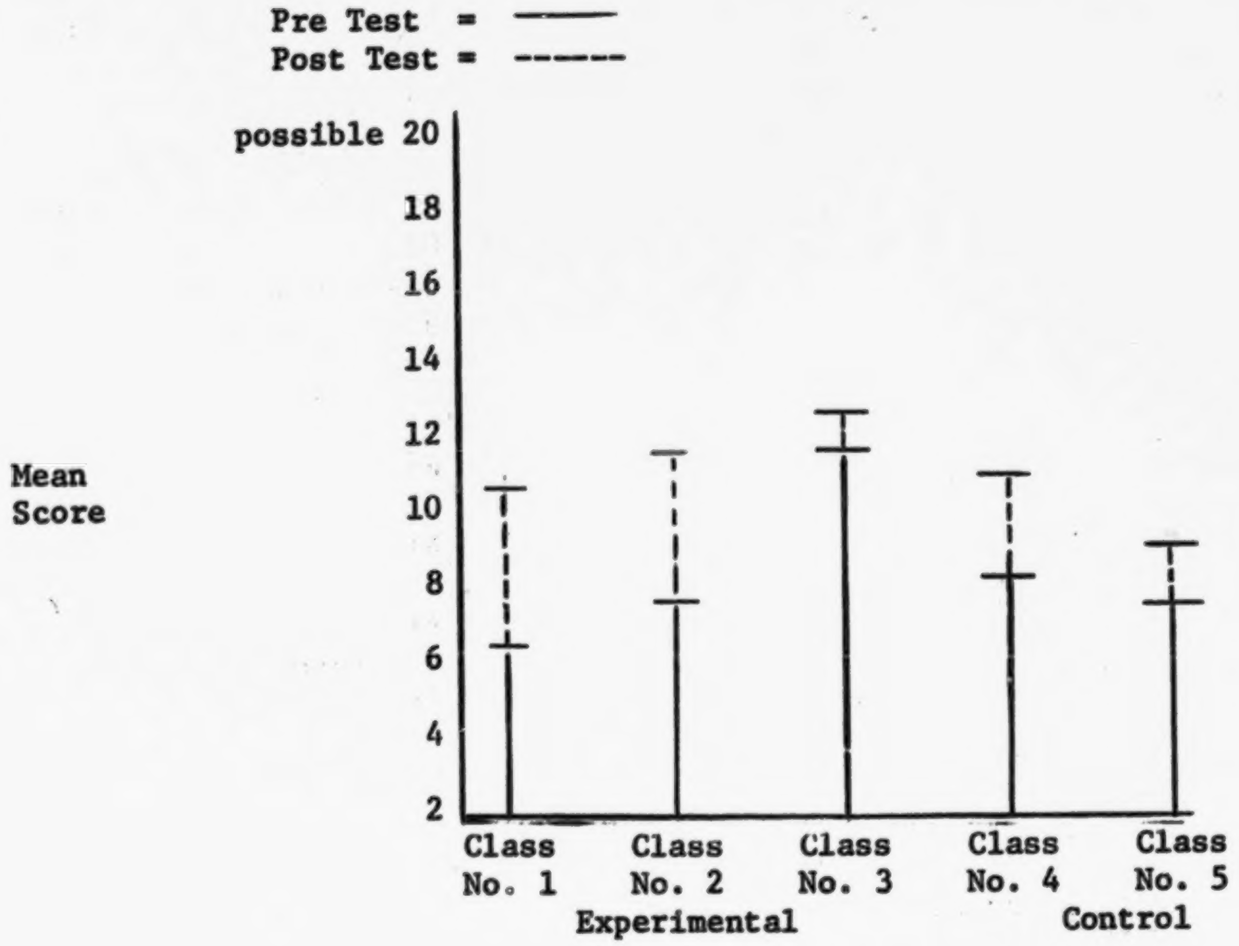


3

ALPHABET - UPPER CASE



ALPHABET - LOWER CASE



HEAD START EVALUATION AND RESEARCH CENTER

TRAINING FOR NUMBER CONCEPT[1]

A.H. Blum

Boston University

ABSTRACT

In order to study the emergence of number concept in preschool middle-class and disadvantaged children, two diverse samples were given training and testing in number conservation. Specifically, determination was made as to whether conservation of number emerges at different times and whether training procedures have differential effects. Subjects were tested under two treatments (manipulation of stimuli) and two conditions (correspondence). Differences were found between groups and as a result of training. Absence of interaction indicates that disadvantaged children are as amenable to training as their middle-class peers.

---

[1] "The research reported herein was performed pursuant to a contract with the Office of Economic Opportunity, Executive Office of the President, Washington, D.C., 20506. The opinions expressed herein are those of the author and should not be construed as representing the opinions or policy of any agency of the United States Government."

TRAINING FOR NUMBER CONCEPT[1]

A.H. Blum[2]

Boston University[3]

The revival of interest in Jean Piaget's theories of child development is directly related to a revival of interest in cognitive development. Specifically it is based on the "need" to test and better understand his theory of the stages and processes involved in intellectual functioning. An important concommitant of the latter are attempts to develop new techniques and methods to facilitate the preparation for and learning of the basic tools and concepts related to specific school subject content.

This study which is concerned with the general area of emergence of number concept, specifically, investigates the attainment of the concrete operation of number conservation[4], and the efficacy of a training procedure to facilitate and "speed-up", age-wise, number conservation. As Piaget suggests, the child "does not first acquire the notion of quantity and then attribute constancy to it; he discovers true quantification only when he is capable of constructing wholes that are preserved". (1965, p. 5)

There have been a number of studies which have attempted to induce conservation or facilitate its emergence in a particular area of the concrete operations stage. Training in addition-subtraction failed to produce conservation of weight and substance (2,4). Although such conservation was induced by training involving "conflict" situations (4), it is not clear what variable accounted for the emergence

---

[1] "The research reported herein was performed pursuant to a contract with the Office of Economic Opportunity, Executive Office of the President, Washington, D.C. 20506. The opinions expressed herein are those of the author and should not be construed as representing the opinions or policy of any agency of the United States Government."

[2] Mrs. V. Carolyn Adcock is the research associate on this project, without whom this study would not have been possible. Our cordial appreciation is acknowledged to Mrs. Joan Needleman of the Wayland, Massachusetts School System, and to Mrs. Elaine T. Pelavin of the San Francisco School Department

[3] Now at the University of Wisconsin--Milwaukee

[4] Broadly stated, a concrete operation is any representational action which is basic to, and integrated in an organized framework of related acts which are utilized by the individual to organize and manipulate the "realities" of the world around him. Conservation is the act of organizing perception so that any given physical dimension (in this case, quantity) is held constant despite any changes of irrelevant conditions, e.g. a given number of items is seen as equal to another set of items identical in number despite any change in spatial arrangement of either set of items.

of conservation or whether the training simply forced the child to "stop and think" about his answer.

On the other hand, counting and addition-subtraction training failed to produce number conservation in a study by Wohlwill and Lowe (1962). Similarly, Wallach, Wall, and Anderson (1967) found addition-subtraction training singularly ineffective in training for number conservation. In this same experiment, however, reversibility training did facilitate number conservation, although this ability did not transfer to conservation of discontinuous substances.

Interestingly, what is repeatedly cited as characteristic of non-conservers, the inability to disregard irrelevant perceptual cues, has not been investigated empirically. Yet, it is precisely this ability which may define conservation or non-conservation, for implicit in a child's attainment of conservation in general and number conservation in particular is the fact that he has stopped using the misleading perceptual cue of "how it looks." This would suggest that training which enables the child to resolve and ignore irrelevant perceptual cues may, in and of itself, facilitate the attainment of number conservation.

Specifically, this study hypothesized that both non-conservers and "transitional" children (Stage 1 and 2 children) could be trained to deal with perceptual confusions so that they could utilize this understanding to disregard irrelevant changes such as spatial rearrangement, thereby maintaining the notion of conservation of number (Stage 3). (Three stages in the attainment of conservation of number are outlined by Piaget: 1) No conservation. In this stage "judgments" involving quantity are purely perceptual, i.e., based solely on appearance." 2) Transitional stage. At this level concrete operations are used "off and on", i.e., conservation of quantity lacks consistency and/or permanence. 3) The stage of conservation number.)

As further assessment of the effectivenss of such a training procedure, two diverse experience groups were compared under differing conditions. The two subject populations were Headstart and Title I children and their "middle-class" age peers. Specifically the two populations were compared to determine 1) whether conservation of number emerges at different times for the two groups and 2) whether the training procedures have differential effects for the two groups.

METHOD

S's were 109 children, 45 "middle-class" youngsters and 64 Headstart and Title I youngsters. Ages ranged from 4;2 to 8;11, representing five age groups: Eights, Sevens, Sixes, Fives and Four-year-olds. Mean ages for the groups were 8;5 (N=21), 7;6 (N=21), 6;6 (N=21), 5;5 (N=38), and 4;9 (N=8).

S's were tested under two experimental treatments for each of two conditions.

The difference between the two experimental treatments was whether manipulations were made with S's or E's set of stimuli; the two conditions were "provoked correspondence" and "spontaneous correspondence."[5]

Only S's who demonstrated understanding of the concepts "more", "same", and "less" and were able to count to nine were tested for conservation. Presence or absence of conservation was determined by the classical Piaget-type pretest described below. All S's making one or more incorrect responses on either the provoked or spontaneous parts of the pretest received the training procedures or served as controls. To avoid making the sessions too long and repetitious, the second half of the pretest (spontaneous) was not administered if S failed to conserve on any part of the first half (provoked).

S's were administered the pretest and training sessions in each type of conservation the first day (pretest and training sessions were separated by an irrelevant "game" involving pictures of cats and dogs). On the second day S's received an additional training for each conservation and following another irrelevant "game" (involving colors and shapes), were post-tested for each conservation. Pre-and post test procedures were identical. With one exception, each session utilized different stimuli; no stimuli were used twice on the same day.

Pre-and Post Testing Procedure: After showing S the stimuli to be used for a particular session (in provoked sessions, S was asked to "put together" one pair of the stimuli; in spontaneous sessions, S chose which color he would like to play with), E made a row directly in front of S with one color or one half of the stimuli. S was then instructed to make a row "just like E's row (S was always given two pieces more than E used to construct his row). If necessary, after S constructed his row, E assisted S in making his row equivalent to E's row. When the rows were equal, E pointed to S's row and asked for example, "Are there more roofs, the same number of roofs or less roofs than there are (pointing to E's row) houses?" and after S's answer, "How can you tell?" Those few subjects who denied equivalence were eliminated from further testing.

E then either extended or condensed one of the rows and asked S again if there were: more, the same number or less pieces in his row and how he could tell. After the row had been returned to its original position and S had answered the questions, the opposite manipulation to what had preceded was made, followed by the row again in its original position.

---

[5] Piaget differentiates between "provoked" correspondence and "spontaneous" correspondence, noting that in the former there is a functional or complementary relationship between two equal numbered sets of objects which facilitates making correspondence and hence equivalence, while spontaneous correspondence simply involves identical sets of equal numbered objects. Provoked correspondence is said to emerge earlier.

Training: The training sessions differed from the above only in the use of "long," "neutral", and "short" spaces (cardboard pieces 2 1/2 inches, 1 1/2 inches and 1/2 inches, respectively) for constructing and manipulating the rows. After S demonstrated that he understood the concepts "longer," "the same length" and "shorter" E constructed his own row using the neutral spacers and had S also use them in the construction of his own row. Before putting other spacers into either row E asked for a prediction, "What is going to happen when I put these in your (my) row?" When an incorrect prediction was made, E said, "But I'm not changing the number, I'm only changing how long the row is." For each manipulation, extension, condensing or return to original, S was asked for his prediction in addition then to his judgment as to the equivalence or nonequivalence of the rows. If correct, S was asked how he could tell. If incorrect, E asked him to count the pieces in each row, then again asked whether the number of pieces in each row was the same. If S still asserted inequality, E said, "let's do this and see what happens," trading--pair by pair--the spacers in the two rows. If S answered the equality/inequality question incorrectly after this manipulation, he was again asked to count the pieces in each row and again was asked whether there were more, the same number or less. Whether correct or incorrect, E then went on to the next manipulation. Each step of this latter procedure was used only if S's previous answer was incorrect.

To serve as controls, nine Headstart and six M-class children who failed to conserve during pretesting performed a number-related task (matching cards with differing numbers of holes to peg formations of the same number) instead of receiving training.

Results

(Since this is an ongoing study the "results" are given in very unsophisticated form and are meant to be purely tentative and suggestive).

In each of the age groups for the two experience groups, children were categorized as conservers (no training), non-conservers and transitional conservers (to be trained) or as untestable (not meeting criteria for sample inclusion).

Table I presents a classification breakdown of the children in the two experience groups for each age group who were not part of the training or control groups.

TABLE I

CLASSIFICATION OF SUBJECTS

	MIDDLE CLASS				HEADSTART/TITLE I				
AGE	8	7	6	5	8	7	6	5	4
CONSERVERS	6	10	7	1	6	2	1	1	
ELIGIBLE FOR TRAINING		3	9	4	9	6		1	
EXPERIMENTAL N=	6	13	16	5	15	8	1	2	
DO NOT MEET SAMPLE INCLUSION CRITERIA			3	2			1	29	8
TOTAL SAMPLE N=	6	13	19	7	15	8	2	31	8

In Table I some interesting differences are found in conservation ability when we compare the two experience groups: at age eight all middle-class children are conservers whereas only 40% of the Headstart children indicate this ability and at age seven 77% of the middle-class children as compared to 25% of the Headstart children can conserve. (Differential N's at the younger ages do not allow for sensible comparisons at this time).

Table I also shows that whereas only two of the seven middle-class five year olds did not meet the sample inclusion criteria (ability to deal with the concepts "more", "same", and "less") 29 of the 31 Headstart five year olds could not handle these terms.

Table II presents the results for trained and control S's in the two experience groups. (It should be remembered that training was given only to those S's who passed criteria and did not have full conservation). Because the N's are small, no attempt is made to analyze the data in terms of age groups but rather the performance of non-conservers and transitional conservers who did not receive training is compared.

TABLE II

TRAINING GROUPS

experimental

		MIDDLE CLASS				HEADSTART/TITLE I			
	AGE	8	7	6	5	8	7	6	5
PRE-TRAINING STAGE	POST-TRAINING STAGE								
NON-CONSERVERS	FULL CONSERVATION		1	2	3	2	1		1
	TRANSITIONAL			1			1		
	NONE		1	1					
TRANSITIONAL CONSERVERS	FULL CONSERVATION					1	1		
	TRANSITIONAL		1						
	NONE								

control

		MIDDLE CLASS				HEADSTART/TITLE I			
	AGE	8	7	6	5	8	7	6	5
PRE-TRAINING STAGE	POST-TRAINING STAGE								
NON-CONSERVERS	FULL CONSERVATION								
	TRANSITIONAL								
	NONE			4	1	4			
TRANSITIONAL CONSERVERS	FULL CONSERVATION						1		
	TRANSITIONAL			1		2	2		
	NONE								

It is quite apparent from Table II that the training sessions were successful in facilitating number conservation. Looking at all ages for both groups together, of those children who initially demonstrated no conservation, ten (71.4%) achieve full conservation on the post-test, two (14.3%) progress to the transitional stage and only two children make no progress. None of the nine non-conservers who were not trained changed their performance in any way.

Table II also indicated that two of the three transitional conservers (again looking across ages and groups) achieve full conservation while one does not change. On the other hand, five of the six transitional subjects not trained do not change in conservation ability while only one demonstrated mastery of conservation in the post-test. Finally, some interesting differences were found between the two experience groups in effectiveness of the training procedure with seven out of the ten middle-class children "changing" and seven out of seven Headstart children showing this "change".

Discussion

The training procedure developed for this study appears to be successful with virtually all ages and is successful with both middle-class and Headstart Title I children. It should be noted that the only trained youngsters, (1 transitional, 2 non-conservers; all middle-class), who failed to change their performance "toward" conservation were the first three subjects trained, raising a possible question of problems of administration.

That the training procedure "works" is important in and of itself since previous efforts have been, by and large, unsuccessful. Additionally, this study points to the generally known fact that "underprivileged" children are "behind" when they enter school but probably more important is the finding that such youngsters are as amenable to training as are their "more privileged" peers. As mathematics is part of school curriculum from the very beginning and conservation of number is basic to a concept of number, a very practical application of this training procedure would be to use it to erase or ease one of the differences that can only snowball into more insurmountable difficulties with time.

REFERENCES

1 Piaget, J. The Child's Conception of Number. New York: Norton, 1965.

2 Smedslund, J. The acquisition of conservation of substance and weight in children: II. External reinforcement of conservation of weight and of the operations of addition and subtraction. Scand. J. Psychology, 1961, 2, 71-84.

3 Smedslund, J. The acquisition of conservation of substance and weight in children: IV. Attempt at extinction of the visual components of the weight concept. Scand. J. Psychology, 1961, 2, 153-155.

4 Smedslund, J. The acquisition of conservation of substance and weight in children: V. Practice in conflict situations without external reinforcement. Scand. J. Psychology, 1961, 2, 156-160.

5 Wallach, Lise, Wall, A. J. and Anderson, Lorna. Number conservation: The roles of reversibility, addition-subtraction, and misleading perceptual cues. Child Development, 1967, 38, 425-442.

6 Wohlwill, J. F.,and Lowe, R.C. Experimental analysis of the development of the conservation of number. Child Development, 1962, 33, 153-167.

HEAD START EVALUATION AND RESEARCH CENTER

A Study of Preferences Among Qualitatively Differing Uncertainties[1]

David Mostofsky

Boston University

ABSTRACT

Head Start children participated in an experiment in which rewards were made available. Regardless of the child's position response (right or left) the probability of reward was always $p=.5$. Discriminative stimuli were made available; one side imperfectly correlated with the subsequent availability of reward, the other perfectly correlated. Preference for the "consistent side" was evidenced. Implications for application as a non-verbal diagnostic and training model are discussed.

---

[1] "The research reported herein was performed pursuant to a contract with the Office of Economic Opportunity, Executive Office of the President, Washington, D.C. 20506. The opinions expressed herein are those of the author and should not be construed as representing the opinions or policy of any agency of the United States Government."

HEAD START EVALUATION AND RESEARCH CENTER

A STUDY OF PREFERENCES AMONG QUALITATIVELY DIFFERING UNCERTAINTIES[1]

David Mostofsky

Boston University

Problems in thinking and concept formation have had a long history in the literature of Psychology. It is only relatively recent, however, that the focus of interest has turned to an adaptation of refined techniques in response to questions of concern such as decision processes in children as a function of stimulus as well as cultural conditions has been observed by many. The Educational Testing Service has for one had more than a casual interest in the nature of such choice responding. A series of papers (representative of the work by Rosenhan (1966a) is evidence of such studies. The Rosenhan type of study has not contented itself with the examination of alternation behavior in the neutral laboratory, but has been equally concerned with the affects of social class and race on responsiveness to the effects of reinforcement (Rosenhan 1966b). The purpose of the present study is also directed to a specific activity of choice and preference. The particular interest in this study is to utilize a non-verbal task which can give evidence of preferences exhibited under conditions of equal satisfaction(reward). Such a demonstration would suggest the ability to control or maintain behavior without explicit administration of a reinforcing agent. Historically, this problem might be though of as beginning with a master's dissertation by Prokasy(1956) and extended more recently by Bower et.al(1966). The model inherent in these investigations may also be seen in a series of studies pursued by Weir (1964, 1965) in which children's preferences were observed independent of reward consequences.

A number of conditions strongly argue for investigating alternative behavioral techniques which might be of general value both for diagnostic and remedial purposes. Principally, such techniques would place minimal emphasis or requirement on the use of verbal behavior. To the extent that debilitations of either a cultural of intellectual nature affect the learning or other performance capabilities of children and adults, the wisdom of using any verbal based investigatory or measurement scheme seems highly questionable. A host of theoretical issues surround this problem of non-verbal alternatives, not the least of which concerns a position formulated most explicitly by B,F. Skinner, viz. that the most complicated classes of human and infrahuman behaviors are subject to the same fundamental laws of control. In dealing with the problems and objectives of learning in children, and adults, the advent of program instruction and teaching machines have tended to demonstrate the efficacy of reinforcement and its associated schedule of dispensation. The data tend to convey an expectancy of success in the modification or maintenance of verbal and other behaviors by recourse to known and existing reinforcement schemes. These procedures and schemes usually differ with respect to parameters such as the magnitude, frequency, and rate of reinforcement availability. Staats and Staats(1963) in their work "Complex Human Behavior" discuss a number of variables which have successfully been explored in conjunction with the study and control of human behavior.

---

[1] "The research reported herein was performed pursuant to a contract with the Office of Economic Opportunity, Executive Office of the President, Washington, D.C. 20506. The opinions expressed herein are those of the author and should not be construed as representing the opinions or policy of any agency of the United States Government."

Outside the laboratory situation, however, the realities of utilizing primary reinforcement would seem constraining features for large-scale implementation. Instrumentation problems aside, certain sub-classes of the subjects, who might otherwise benefit most from the favorable features of these procedures, may well suffer from unanticipated problems, such as the utility of the reinforcement; i.e. the subjective utility of the reinforcement as perceived by the subject. Such a hypothesis was advocated by Rosenhan in a series of studies executed under the aegis of ETS. Their studies suggested the instability of reinforcement utilities, and that the perceived value of the reinforcement differs as a function of racial group membership. It would appear desirable, therefore, to (1) capitalize on the desirable qualities of the reinforcement and operant approach, (2) incorporate the technology of the experimental analysis of behavior in investigating not only learning behaviors, but the less articulatable qualities which surround the learning situation, and (3) to bypass the liabilities of utility, satiation, etc. which accompany the use of p primary reinforcement. The present study represents a modest attempt to satisfy these criteria.

The theme of this study may be simply described as measuring non-verbally the preference among alternatives whose net probability of payoff is the same. Weir(1964) had investigated a situation where children were able to manipulate one of two plunges-left or right. The left plunger was programmed so as to dispense reinforcement alternately i.e. 50% of the time (every other response), while the right plunger was also programmed to dispense reinforcement 50% of the time, however on a random basis. In general, the left plunger was preferred. Subsequent extensions of this design enabled Weir to program the right plunger for values other than 50%. Weir interpreted his experiment as demonstrating preferences under the conditions of consistency (alternate) versus non-consistency (random), since children indicated a preference for the "consistent" side. There are several difficulties in this kind of interpretation: (1) to identify the preference as between the consistence versus non-consistence, does not offer any additional explanatory value. The terminology or concept of consistency as used essentially describes the technique alone. While it may be a convenient label for the experimenter to differentiate the plungers with respect to the rules with which reinforcement is made available, it has no other embellishingor connotative properties. It certainly contributes nothing to an account of the subject's behavior. It only accounts for the behavior of the machinery. (2) The Weir experiment, when looked at closely, is nothing more than a comparison of components in a concurrent schedule. The alternating schedule (in which every other trial pays off) is a fixed ratio of 2 (FR2), whereas the random schedule is a variant of the VR or VI schedule. Weir's data, therefore, are consistent with the comparisons of data from incompatable concurrent schedules as described by Catania(1966). It does not, however, demonstrate control by any feature other than the schedule.

A different paradigm was, therefore, adopted in the present study. The approach derives from a master's thesis by Prokasy who used rats in a T maze situation. The legs of the T were varied such that the right leg, painted striated, always led to a delay chamber which was also striated. Following a delay period in the striated chamber the animal was released to the end of the arm in which reinforcement was either made available or not made available, with a probability of .5. If the animal chose the left arm of the T maze, he entered a delay chamber which was either striated or solid color. If it was striated, release from the chamber was constantly followed by reinforcement. If it was not striated, release was never followed by reinforcement i.e., CRF versus extinction. Again, the probability of CRF or extinction was .5.

The net expected probability of payoff to either side of the T maze was, therefore, the same. The animal's preference for one arm over the other would reflect his preference for a payoff consistently associated with some discriminative property of the delay chamber rather than a payoff not consistently associated with some discriminative property of the delay chamber. Prokasy found a preference for "the consistent side." His interpretation relies heavily on the hypothesized optimization of anticipatory salivating behavior which the animal undertakes in the delay chamber and which facilitates the terminal consummatory response. Such a preference, if found with humans, would not lend itself to the anticipatory salivating explanation, although it very well might lend itself to an explanation of added reinforcement in a form of conditioned reinforcement provided by the association of the discriminative stimulus with the payoff. It is this model which has been incorporated in the present study. The data would enable a general statement of preference (the control of performance) between responses of different topologies having equally objective probabilities of payoff by subjects for whom learning and cultural deprivations may differ. This differs from the Weir experiment in that the reinforcement schedule is identical in both response models. Consistency may thus be meaningfully ascribed to the situation in relation to the predictability of reward subsequent to the onset of a discriminative stimulus.

Procedure

Subjects were brought to a room (which was part of a two-room suite) in which one wall was a one-way observation mirror. The subject's room housed a response console (see figure 1). The subject's working area on the console contained two buttons (Microswitch#2C206) which when depressed actuated a double pole-switch. The buttons were 5" apart. Each unit could be illuminated independent of either the subject's responses or the switching function. Facing the subject was a screen which deflected pennies which were dispensed from a Gerbrand feeder mounted in the rear of the console. Above the feeder-shield was mounted a resettable six-digit counter. Control equipment was housed in the adjacent experimenter's room. The control equipment was programmed such that continuous operation of the right button caused it to become illuminated yellow on the average of every 20 seconds. Once a switch became illuminated all other functions on the console remained inoperative until that same lit key was depressed again. The subject, therefore, had to respond to the yellow button (i.e. the illuminated manipulandum) for any functional change to take place for him. Responding to the yellow button led to a dispensing of reinforcement on a random average of 50% of the time. The subject's reinforcement consisted either of an increase in the counter, a dispensing of a penny, or both. The button light would then go off and the session continued. In the case of children the use of the penny was defined not for its own utility but rather served as a token. Prior to the session the subject was invited to a room containing a toybin from which he selected a toy of his choice. He was then told he would be playing a game for which he might earn pennies and each penny was to be placed in a bank which was a transparent jar. When the jar was filled the toy of the child's choice would then be given to him. Both the toy and the jar were constantly in the subject's view during the course of the experiment (procedure of Staats and Staats).

With respect to the right button, therefore, on the average it would become lit, the probability of reinforcement following yellow, equal to .5. Continual operation of the left button would occasionally (VI of 20) also turn color. However, 50% of the time it would turn red, 50% of the time, green. The consequence of operating the left button once it was lit was perfectly correlated with its color. If the button was red, its operation would simply lead to the termination of the color and no reinforcement made available. If the button was green, its operation would lead to a dispensing of reinforcement followed by termination of the color and the session continued. All responses and latencies were recorded. Response rate was recorded on standard cumulative recorders. The subjects for the experiment were recruited from an available Headstart center in operation in Revere, Massachusetts. They were transported in groups of three from Revere to the experimental chambers at our laboratory and were individually tested for the duration of the experiment. Other subjects were selected for the experiment from among those participating in on-campus programs. The subjects included adults as well as children. All subjects were first tested for color blindness using the Dvorin Pseudoisochromatic Charts. The analyses of interest with respect to the data, concerns the preference of the left red/green button over the yellow button, i.e., the consistent versus the non-consistent side with respect to a) total number of responses, b) the relative frequency of responding, c) the rate of responding and d) the latency of responses, respectively. The experiment is continuing for several reasons: (1) the sample size precludes the use of powerful tests and, therefore, inadequate for any definitive conclusions, (2) certain control conditions are yet to be run, among them extinction following preference behavior and the use of yellow-blue each with a 25% reinforcement compared to the red-green (as a counterbalance for the possible novelty effects).

The results give evidence that a) children will prefer a "consistent" reward situation to a reward uncertainty situation. This preference was exhibited to some degree by each child tested. The data taken collectively are in agreement, i.e. response rate, total response output, and response latency. b) Adults, while seeking to arrive at an optimal "strategy" for maximizing payoff, also behave under the control of the control of the uncertainties. c) The individual differences and variability exhibited in the extent of preference can be minimized - and the preference correspondingly accentuated - by introducing a change-over delay. We have done this manually and plan in succeeding experiments to build it in as a feature of our logic and control instrumentation.

These findings strongly argue for sober reconsideration of the effectively controlling stimuli in an applied learning situation. While not deprecating the force of reward per se, the data clearly imply that much learning can be efficiently directed by the manipulation of other environmental features - some of which may be considered to have acquired secondary or conditioned reinforcement properties. It would seem imperative to attempt a translation of these findings for implementation in the classroom. The need is even greater where mere increases in reward is contraindicated. The experiments will be continued and will involve larger numbers of experimental units for adults, normal children, retardates. In addition, supportive experimentation will be concurrently pusued i.e.

1) extinction following stabilized behavior (using existing schedule)

2) extinction following stabilized behavior (using change-over delay)

3) forced trials to each manipulandum condition

4) red-green us yellow-blue (instead of r-g us y-y)

The conduct of such laboratory investigations surely constitutes a necessary prerequisite for effective large scale adaptation and for increasing the effectiveness of educational practices in the real and applied world.

The data analyzed thusfar suggests that the predictions which would have been expected on the basis of reinforcement theory are substantiated in these studies. As such, its suggested relevance to the applied situation becomes quite substantial if the reliability of these findings are justified. That is, it does suggest that the use of secondary reinforcement may be an applicable reinforcer for Headstart programs. If this is the case, the failings which are encountered with primary inforcement situations might be avoided and would not be sensitive to individual differences with respect to utility, satiation, or other factors related to racial and ethnic composition. Furthermore, it strongly suggests the advisability of paying closer heed to whatever reinforcement qualities attend the consistency of reinforcement schedules. It further suggests that the reinforcing advantages of discriminative stimuli should not be overlooked even when primary reinforcement is used. Perhaps a behavioral analysis of "consistency" may yet be profitable.

REFERENCES

Catania, A.C., Concurrent Operants, in W. Honig(ed) Operant Behavior: Areas of Research and Application, 1966 Appleton Century, N.Y.

Prokasy, W.F., The Acquisition of Observing Responses in the absence of differential, external reinforcement, Journal of Comparative Physiological Psychology 1956 49 page 131-134.

Rosenhan, D.,(a), ETS Research Bulletin, 66-1, 1966

Rosenhan, D.,(b), Double Alternation in Children's Binary Choice Responses, Psychonomic Science 1966,4No. 12 page 431-432.

Rosenhan, D.L., The Effects of Social Class and Race on Responsiveness to approval and disapproval, Journal of Personality and Social Psychology 1966,4No. 3 page 253-259.

Staats, A and Staats C., Complex Human Behavior, 1963 Holt-Rinehart, N.Y.

Weir, M.W., Children's Behavior in a Two Choice Task as a function of Patterned Reinforcement following forced Choice, Journal of Experimental Child Psychology, 2 page 85-91, 1965.

Weir, M.W., The Effect Patterned Partial Reinforcement on Children's Performance in a Two Choice Task. Child Development, 1964 35 page 257-264.

OTHER PROJECTS

HEAD START EVALUATION AND RESEARCH CENTER

THE UTILIZATION OF NON-PROFESSIONAL INTERVIEWERS IN THE NEW ENGLAND AND MISSISSIPPI SAMPLES BY THE BOSTON UNIVERSITY HEAD START EVALUATION AND RESEARCH PROGRAM, 1966-1967[1]

Suzanne Clay

Boston University

ABSTRACT

This report is a description of the steps taken by the Boston University Head Start Evaluation and Research Center in cooperation with participating Head Start programs in 1966-1967 to utilize non-professional interviewers, 24 mothers of Head Starters in the collection of data about 191 samples families in New England and Bolivar County, Mississippi. It includes a discussion of the characteristics of the non-professionals hired, the nature of the training given, and some assessment of their performance in the Evaluation Project.

---

[1] "The research reported herein was performed pursuant to a contract with the Office of Economic Opportunity, Executive Office of the President, Washington, D.C., 20506. The opinions expressed herein are those of the author and should not be construed as representing the opinions or policy of any agency of the United States Government."

THE UTILIZATION OF NON-PROFESSIONAL INTERVIEWERS IN THE NEW ENGLAND AND MISSISSIPPI SAMPLES BY THE BOSTON UNIVERSITY HEAD START EVALUATION AND RESEARCH PROGRAM, 1966-1967[1]

Suzanne Clay

Boston University

This report presents the highlights of the descriptive material and interviews emergent from the Boston University Head Start Evaluation and Research parent interviewing project conducted among sample families living in the New England area and Bolivar County, Mississippi. This project, developed during the period of March through August 1967 was a part of the national evaluation of Head Start.

Because the project involved the hiring of 24 non-professionals or "community interviewers" who were Head Start mothers with prior experiences working or volunteering in Center activities, it appears that the organizational steps used constitute "research firsts" in several ways and, thus, warrant their reporting. The emperical data represents the responses of 191 parents in the New England and Mississippi samples to a series of questions contained in the Parent Interview.

Organization of the Report

The report is organized into four sections:

Section I - Development of the Parent Interviewing Project
Section II - Description of the Training
Section III - The Mississippi Interviewing Project
Section IV - An Assessment of What Was Accomplished in the New England and Mississippi Project

Sections I through IV are more descriptive than interpretive, simply because it seems that to get a full picture of the characteristics of the non-professionals hired, the role they played, and some assessment of their performance in the evaluation requires an account of how the project was developed from the outset.

Section III focuses specifically on the procedures employed in Bolivar County, Mississippi, in an attempt to describe some of the unique problems encountered in this area.

---

[1] "The research reported herein was performed pursuant to a contract with the Office of Economic Opportunity, Executive Office of the President, Washington, D.C., 20506. The opinions expressed herein are those of the author and should not be construed as representing the opinions or policy of any agency of the United States Government."

SECTION I - DEVELOPMENT OF THE PARENT INTERVIEWING PROJECT

A major project of the 1966-1967 evaluation of Head Start children in New England and Bolivar County, Mississippi, was the utilization of non-professional community interviewers in the collection of information about the attitudes and reactions of parents during an interview. Our decision to recruit and train community people resulted from requests by several New England Head Start Directors that the Boston University Head Start Evaluation and Research Center make a commitment to involve parents of Head Start children to the fullest extent in meaningful and productive roles within the evaluation and research design. A secondary factor supporting the suggestion was our interest in demonstrating that a team of parents who were knowledgeable about the various communities in which the Head Start evaluation samples were located and who had training in interviewing skills and on-the-job supervision from E&R Center Staff would be able to elicit cooperation and reliable responses from parents of Head Start children.

Accordingly, during the months of March and April 1967 the research associate assigned to develop this phase of the evaluation met with Directors of sample Head Start Centers in New England and requested their assistance in developing a feasible plan of operation. Among the many questions posed were the following: What impact will the activities of teams of community interviewers have upon ongoing Head Start programs? Will the parents resent having neighbors in some instances, asking probing questions about their children's behavior at home? Will the parents, many of whom are receiving welfare assistance, be confused about the Head Start evaluation project to the extent that they will identify the community interviewer as an employee of a welfare agency? Will there be avenues of communication built into the project so that professional Head Start social service workers can share in planning methods to be used? What financial arrangements will be made to make this a reasonable task for the community interviewer? Will the position be advertised widely so that the better applicants may be chosen?

Social Workers Seminar

Obviously it was apparent that given the design of utilizing non-professionals in the communities as parent interviewers, the fulfillment of our goal depended upon the close cooperation and joint planning of the E&R staff and personnel in each of the sample centers. Head Start Directors in New England sample centers were asked to appoint two representatives from among their total staff of social service workers, neighborhood case aides, social service directors and chiefs of parent programs, to attend a series of four seminar planning meetings at Boston University during the months of May and June 1967. The focus of the seminar would be to explore many questions that had been raised and to further establish a forum whereby these staffs in the sample Head Start Centers could be apprised more fully of the role of evaluation and research and the progress of the parent interviewing project.

These Spring afternoon meetings proved to be successful in accomplishing our initial goal to plan for recruitment, hiring and supervision of community interviewers. The secondary goal of facilitating new channels of communication between Head Start social service staffs throughout New England was met to a surprising degree. It was observed at these meetings that Head Start personnel who were employed in comparable roles in various communities needed a medium to exchange ideas as well as to air frustrating experiences in order to assess their own situation in clearer perspective. That personnel in rural communities were unaware of the social blight found in inner city ghettoes was apparent. As the group became more knowledgeable about differences between Head Start programs operating in urban city and rural suburban areas and more comfortable in verbalizing experiences, many suggestions for "successful remedies and techniques" in building a strong parent involvement program within Head Start programs were shared.

To facilitate discussion of some of the issues facing poverty workers in the area of social services, the group was shown a film at the third session that focused upon the frustrations, aspirations and fears of America's poor. Tapes of interviews were played back to this group of professional Head Start workers for their reactions.

Qualifications of the Community Interviewers

Specifically, the results that issued from this communication with social service staffs in sample Head Start programs throughout the New England area was the result of joint planning of a design that mapped out what qualifications the community interviewers should have as well as suggestions to E&R regarding topics that should be covered in training sessions. Briefly, the design agreed upon was one involving 12 parents of Head Start children who must have had prior experience working in some phase of the program as a volunteer worker, classroom aide, neighborhood case worker, member of the Head Start policy advisory board or in comparable roles requiring their active participation.

They would be required to attend a series of weekly training and orientation sessions to be held at Boston University throughout a three-month period. Each community interviewer would be paid for attending orientation training sessions, completing satisfactorily assigned parent interviews and personal transportation costs incurred in making visits to homes. Each of the sample Head Start Directors in New England was asked to submit recommendations for the parent representatives to be interviewed for the position of community interviewer within each community. Furthermore, it was agreed that the E&R staff and community interviewers would cooperate closely with Head Start social service staffs in making contacts with those families who were already involved in referral services from the local welfare agencies.

Who Was Hired

By May 1967 we had received the letters of recommendation for 12 mothers from the Boston ABCD, Brockton Self-Help, Incorporated, Hartford Child Development Program and the Cambridge Head Start programs. After screening and approving applicants the E&R training staff conducted a general orientation meeting at Boston University in May, 1967. The mothers had backgrounds that appeared to capture the experiences needed. Two parents had been employed as teacher aides for two years in Cambridge Head Start classes; two were currently serving terms of office as President and First Vice-President of the Hartford Child Development program in addition to sharing with the total group roles as mothers - active and vocal in their local P.T.A.'s and church groups. Although the qualifications sent to Head Start Directors did not state a preference for the better educated, it seemed that their ability to function successfully in community organizations, relevant work experience and alertness to the reactions of others correlated highly with their ability to function as community interviewers.

One mother was Spanish-speaking, and it was quite apparent that her bilingual ability was an important factor for communicating with Puerto Rican and Portuguese parents. The overall incidence of totally Spanish-speaking parents was small enough throughout the New England sample to be handled by enlisting volunteer services of Spanish-speaking students in local colleges when the need arose.

About one half of the community interviewers expressed the desire to get further education and job training whether this was completion of high school credits or admission in college courses during the evenings and weekends. At least two from this group have become involved, subsequent to their employment as E&R interviewers, as employed workers within an urban recreation poverty project, a job that involves them in working directly with parents and their children. The two former teacher aides have since returned to Head Start classes as aides on either a paid or volunteer basis.

Methods of Publicizing in the Sample Communities

A recommendation made by sample Head Start Centers was that the E&R Center should accept suggestions from the Directors as to how the project should be publicized in the communities. The phase of publicizing the purpose of the parent interview, consequently, was developed according to the methods felt to be the best by the staffs in the sample Centers. For this reason, there was much variation in approaches made to parents from community to community. Particularly at this juncture of the planning an effort was made to carry out the wishes of participating staffs in approaching their parents. The Hartford Child Development Program used a combination of letters to parents from the E&R Project Director and articles in their own newsletter as a vehicle of communication. Another center with a relatively small number of children, the Saint James' Ecumenical Center in Roxbury, Massachusetts, elected to inform their parents via a discussion of the program at one of the weekly parent meetings in the center. The Director of the Cambridge Head Start program felt that their social workers could be very helpful in mentioning E&R involvement in the community during their daily visits to families.

With the exception of the Head Start Center mentioned previously, all of the others supported our wish to mail letters to the sample families before the community interviewers made a first call to homes. The form letter included the purpose of the interview, supplied the parents with a brief statement regarding the background of the interviewer who would be calling in the near future, and introduced the interviewer by name. This technique was evaluated later by both the team of community interviewers and the Head Start Directors as a successful one in building a friendly relationship with individual families.

Finally, contact was made with local police departments and schools in several communities by Head Start Directors to further insure a welcome reception to interviewers as they sought families in public housing developments and apartment buildings. Each interviewer also carried a letter of identification with her to present to the respondent before the interview was begun.

SECTION II - DESCRIPTION OF THE TRAINING

Reviewing the nature of the training sessions, I believe training of non-professionals, with few exceptions, was minimal. The 12 parents in New England received training and supervision over a three-month period in a total of five group meetings and sporadic telephone conversations with the E&R training staff. Two trainers, one an E&R research associate and the other a staff social worker, led the training sessions along the lines of demonstrating how the experiences and problems met by the community interviewers could be handled or resolved to their advantage. Generally, the approach centered around explaining the connection between the E&R design for educational research and the ongoing Head Start programs which the children attended; clarifying to the parent's satisfaction that the interviewer was not a welfare worker seeking to extract highly personal information about the family; and developing interviewing skills among the interviewers so that they could encourage parents to elaborate upon responses or could make probes in a non-offensive manner. The last item was the focus of two of the training sessions in which interviewers elected to role-play specific situations that had challenged their skills, and to discuss how the interviewer should handle the interaction. Several tapes of these sessions were made and proved useful in focusing the group's attention on how they were developing in their ability to interpret the verbal and non-verbal behavior of parents.

Training for the three Hartford community interviewers had to be condensed since they were at great distances from the majority of the community workers located in or around Boston. However, by involving the Chief of Parent Programs from the Hartford Child Development Program (a professional social worker by training as well as a sensitive individual in her relationship with the non-profession) the supervision of these interviewers was also attended to.

Use of "Elastic Group Process"

The key to the training of the community interviewers in New England was the liberal use of what we may term "elastic group process." As the name implies, the object of group process is to accomplish the teaching of specific skills in addition to attending to a system of personal needs within the group. While the interviewers were in the training session, the focus of the discussion was on what had occurred while they visited sample families. It was interesting that in order to define their role, they raised questions about many of the evaluation instruments that were being used within sample Head Start classes. Discussions touched upon the need for testing Head Start children as well as the broad area of educating the disadvantaged child.

Throughout, an attempt was made to maintain sessions as informal forums for airing the problems encountered and exploring approaches to the situation that would produce a sharpening of their perception of interactions.

SECTION III - THE MISSISSIPPI INTERVIEWING PROJECT

Bolivar County, Mississippi, is the location of two Head Start programs involved in the 1966-1967 evaluation. The two Mississippi programs, CAP and ACBC, were located in the Delta region. Families are scattered across wide expanses of rich farmlands in primitive, substandard living conditions. The following is extracted from the narrative account of one community interviewer in this area and captures the realities of day-to-day existence in this county for the majority of E&R sample families:

> It all began the first day of my interviewing in the Benoit district. There were several families in the same neighborhood who were in unlivable conditions--what I mean is poor housing, no water, half-dressed and dirty children. I'm sure the dirt came from lack of water, which they have to haul many miles away by tractor and trailer if the landowner grants permission to use them.
>
> There is little or no work for farm families this season. Fathers can work if they are machine operators; no work for mothers only in some places....They flab weed cotton this season.
>
> Most of the families depend on food stamps since commodity has been out. Families enjoy shopping with food stamps rather than receiving commodity. Although there are some that can't afford food stamps at the times they are due, due to the lack of work.
>
> Many parents asked why they were not in the interview. I explained it to them, They wished they could have been, in hope they could get some aid. I hope conditions will change for the needy in the near future around here.

In this area where Head Start programs have taken a "first step" in getting parents involved in the education of their children, the Head Start program is much more than an educational venture. For a great majority of families in the county it has meant employment, clothes, food and shelter. However, given the existing conditions of the homes visited, it was obvious that there must be a massive attack in the War on Poverty aimed at homes in Mississippi where Head Start children live. Until services providing for physical, nutritional, and social welfare needs are injected into this area, it is doubtful that many parents will be able to address their energies or interests to maintaining a home environment that will augment, or at least, not wreck whatever gains the Head Start child has made.

The model for recruitment, training and on-the-job supervision of 12 mothers hired to do interviewing of the sample parents in the two Mississippi Head Start programs was similar to that developed in New England. However, the great distances between the families of Head Start children and the lack of any form of public transportation for interviewers necessitated that we alter the approach in several respects. It was arranged with the two programs that the team of 12 community interviewers ride on the buses or in the carpools in the mornings or afternoons when the children were being transported to and from centers.

Training and supervision of the field work of the 12 interviewers were the responsibility of the E&R research associate who lived in the region. It would have been extremely difficult to complete the Mississippi interviews given the design adopted without her involvement in the program. That she assumed the responsibility during evenings and weekends to offer individual help to the interviewers in interpreting what the instrument was trying to achieve was a major factor in "getting the job done." Because the interviewers lived in small towns spread over a two-county area, this necessitated that the trainer organize regional training sessions as well as driving interviewers to distant homes.

SECTION IV - AN ASSESSMENT OF WHAT WAS ACCOMPLISHED IN THE NEW ENGLAND AND MISSISSIPPI PROJECT

In general, I believe the community interviewers in New England and Mississippi performed their jobs well. Observations of group sessions indicated that they showed a high degree of perception and insight into problems found within the families and could articulate alternative methods of dealing with them. The following, an excerpt from one of the taped training sessions for the New England interviewers, reflects this quite well:

CLAY: "All right, Ruth, What about your worse case?"

WARD: "This is Gene X. His mother was a million miles away during the entire interview. I knocked at the door. I told her who I was. I had called her on the phone earlier; and, she said, 'Come in,' and walked ahead of me. It was like a role-playing situation. . . . It was real hectic.

(Laughter from the group)

WARD: "She screamed continuously at her children; and, she said to me: 'Would you like a cup of coffee?' And before I had a chance to answer, she said, 'Why don't we sit down?'. . . . Then, trying to interview her was so hard. She didn't look at me. She walked around and she hit one of the kids.
"I smiled sweetly and I talked to Gene about Head Start."
(Pause)

CLAY: "What was the mother doing?"

WARD: "I ignored Mother."
(Pause)
"Well. . . you know, I think the mother resented it when I was talking to him. Maybe she realized I was ignoring her and she said, 'How many more questions are you going to ask me?' Then she came right back and realized I was going to watch her.

CLAY: "Well, how did the interview go?"

WARD: "She answered my questions; but it just took her about 5 or 10 minutes to answer one because she wasn't paying attention to me.

ANOTHER COMMUNITY INTERVIEWER
#1: In other words, she was doing housework?"

WARD: "No, she was just sitting there. She had coffee, had a cigarette, played with the cat, screamed at the kids. . .It was beautiful watching her because I couldn't believe that a mother was actually acting this way. It was a good experience."

CLAY: "Ruth, do you have any idea why she responded to you like that or to the situation? . . . Had she ever been interviewed before?"

WARD: "No, she said she didn't even know the name of the Head Start social worker; and I interviewed about three mothers of the same class and they all knew. They had all seen dental people; they all had seen social workers; and, they all saw nurses. . . . She's never seen anybody."

ANOTHER COMMUNITY INTERVIEWER #2: "Had she ever been to the class? Or from the interview do you have this information with you?"

WARD: "Yes, she took him to school. That's it! That's as far as she went, as far as the classroom is concerned. (Pauses to review interview form). . . To that question: 'If you've not been into the Head Start Class would you like to go?"

COMMUNITY INTERVIEWER #2: "Well, how did she answer?"

WARD: "I asked her that and she said, 'Um?' (Comments among the group) And I asked her if she had noticed any changes in Gene, and she said, 'Yes, he caused less trouble at home, which I appreciate.' And she said he speaks better, he has self-confidence, gets along better with other children....and he's able to do things on his own. From what I gathered though, her other two kids have been receiving help from the Mental Health Clinic. (Pause) . . .I tell you, those two are strictly working on their own 'cause she didn't help them one bit."

CLAY: "Do you think, Ruth, in terms of what you observed in the homes and her attitude during the interview, that she just didn't care about her children sincerely?"

WARD: "That she just didn't care. . .This was my first opinion. . .But, you know, now I don't know. . . . . .She was a very nervous person. She's just a nervous wreck! . . .These are two active little boys who are all over on the chairs, tables, everything! It works on her nerves, I guess. . .She just needs help. I don't mean that she's a mental case or crazy, but needs to calm down, maybe a tranquillizer. I think Gene is her youngest. Maybe she could use a part-time job to get away from the house."

Again, within the comments of the Mississippi interviewers, it is possible to see similar insight into the problems facing the deprived families:

> During my tour through the community while I was interviewing the families, I found families living in unbelievably bad-conditioned homes. Too many children in small towns, such as 16 children living in a three-room house and families who have to haul water many miles away from home are all around.
>
> I feel that mothers would do an overall better job in managing their families if they had better facilities to do it with. Some homes we passed by were in fair condition and others weren't. I think this comes about where in one family there are more workers and less children. In others there seem to be less workers and much more children to provide for.

> In some cases only the women are given something to do and the men and children are left out. If families in our country were given more work to do, I believe things would be in a much better condition. People around here don't even have beds for their children to sleep in, not nearly enough food for them to eat. . .It is a miracle how these little ones even exist.

While these 24 community interviewers were selected on the basis of their previous involvement in the Head Start program and their ability to relate well to others, the training staff felt that closer supervision and more intensive training could yield better outcomes in fulfilling the commitment our E&R Center has made involving non-professionals in meaningful, dignified and worthwhile ways within the Head Start evaluation and research activities.

Several of the questions raised by Head Start Directors and their staffs in New England and Mississippi touch upon providing for the wisest utilization of non-professionals within educational research activities, creating programs within local universities which offer college credits for non-professionals employed in various aspects of the poverty program, and exploring the effects of using non-professionals in jobs traditionally held to be the exclusive province of social workers, teachers and researchers. The exploration of many of these issues lies before our E&R Center within the coming year.

HEAD START EVALUATION AND RESEARCH CENTER

TEACHER SEMINAR[1]

Sandra Alexanian

Boston University

ABSTRACT

The person most often bypassed when describing Headstart is the teacher. Teachers are often told what they are doing - but seldom asked why. This motivated the Evaluation and Research Center at Boston University to ask Headstart teachers to be consultants concerning the description of their program.

Eight Headstart teachers met for seven consecutive two-hour taped sessions. The following areas were discussed at length:

1. Administrative structure of the agency in which they worked.
2. Teacher training prior to working and in-service training.
3. Supervision and support given to teachers.
4. Conditions under which teachers work, and the anxieties generated by them.
5. Assumptions made concerning the needs of children--and the gap between needs and the program offered.
6. The public schools and lack of communication with them.

The seminar was, for most, the first opportunity to meet as a body and discuss mutual problems and needs.

---

[1] "The research reported herein was performed pursuant to a contract with the Office of Economic Opportunity, Executive Office of the President, Washington, D.C., 20506. The opinions expressed herein are those of the author and should not be construed as representing the opinions or policy of any agency of the United States Government."

TEACHER SEMINAR[1]

Sandra Alexanian

Boston University

The teacher seminar was initiated for the purpose of providing teachers with an opportunity to asses their experiences while working in Headstart. Participants came from many types of Headstart agencies (school dominated, rural, urban, single purpose, multi-purpose). They were asked to participate as consultants to the Evaluation & Research Center--the contribution was made with obvious enthusiasm (and relief) by these teachers. They are most often the last person to be asked to contribute to a research project that inevitably describes them. It was a long journey from cathartic discussion to a realistic relating of facts and issues.

The eight participants[2] began to mutually explore the physical and political surroundings in which they found themselves. The initial discussion grew from a natural curiosity about each other; however, it was later clear that there was an absence of opportunity for Headstart teachers to meet outside their own agency. All were amazed at the diversity of "conditions"[3] in which they worked. Each session was tape recorded. The seminar leader[4] in cooperation with Evaluation & Research Center staff members had developed topic areas to be explored. The seminar was held once each week, at Boston University, for seven consecutive weeks. For her services, each teacher was paid as a consultant. Each participant was the "head teacher" in her classroom, a title assigned by the particular Headstart agency for which she worked, since they were responsible only for their individual classrooms.

Headstart Teachers: How do they view their jobs? Although there were a variety of views uttered during the seminars, there was also a striking unanimity of feelings shared. In general, Headstart teachers said their job was to "help deprived and handicapped children." Since defining this aim was so important, the group spent a great deal of time on it. Every teacher made lengthy statements about the administration each time she attempted to describe her position. All felt there were gross inadequacies in the administration of Headstart programs, and that these inadequacies were stumbling blocks in their ability to function effectively as teachers. In many cases, the most serious obstacle to working towards a program goal were the demands made on teacher time outside the classroom. These outside demands included home visits, business meetings, and smoothing out relations with agencies in which their classrooms were housed. One teacher stated that she felt like a "revolutionary" in agency meetings rather than in the classroom. In some instances, the administrative problems of Headstart center were so overwhelming that the very survival of the program was the all important focus. This left the teachers with the conclusion that survival of programs consumed more of their energies than the education of children. Many felt that program survival was an area in which they had no training. A number of agencies were cited as being responsible for the inefficient progress of the child development, Headstart program; among them were public schools, police, government funding agencies, city politicians and welfare agencies. Headstart teachers saw themselves as instructors of preschool children who presently were handicapped, or would be handicapped in the future. Yet the time that was needed to deal with these "handicaps" was often encrouched upon by administrators and community leaders.

---

[1] "The research reported herein was performed pursuant to a contract with the Office of Economic Opportunity, Executive Office of the President, Washington, D.C., 20506. The opinions expressed herein are those of the author and should not be construed as representing the opinions or policy of any agency of the United States Government."

[2] Linda Filiurin, Gail Kursteiner, Dorothy Latham, Edith Martin, Gail O'Connell, Judith Prymack, Mary Rosen, Caroline Wyatt, asst. to seminar director Wilma Snowdon.

[3] Quotation marks are in reference to remarks made on the seminar tapes by participants.

[4] Sandra Alexanian

Being a Headstart teacher can be described from many viewpoints. Income is one aspect that was explored. No teacher in the seminar had a written contract, and there seemed to be little relationship between hours and salary. Within the group where teachers who worked five half days with no home visits, salaried at approximately the same level as teachers with daily morning and afternoon sessions and total parent coverage. There was no financial compensation for many hours of required work undertaken outside the classroom. No teacher has a guarantee of a job for the coming year, nor, for that matter, for the next month. Many were unsure of who their administrators would be, since at present they were "under fire"[5] Program at the administrative level is a major problem.

Experience: Does it influence program adequacy?

The competance of this group of teachers may be unique since there is a large employment pool in Massachusetts from which to recruit. Seven of the eight participants had earned at least one degree.[6] Six of the seven degrees were in education. One teacher held a masters degree. Approximately half the teachers had worked in related educational areas (retardation, day care, settlement house) while the remainder had not worked in the area of sociogenously handicapped children.

Job orientation varied tremendously. Some teachers received a one week orientation and still others received an eight week training program, while one teacher had been hired the week prior to class opening. Most participants felt their orientation programs, though helpful in the areas of songs, stories, games, etc., were naive in the area of social problems. Teachers know before hand that their classes contained problems, but the concepts of "cultural deprivation," "alienated communities," "unrealistic curriculum expectation" needed discussion and demonstration in order to translate these concepts into concrete, educational activities.

Some agencies successfully utilized experienced classroom teachers in their in-service training programs, preschool courses for college credit, staff counselling and support. Unfortunately, much of the "so-called in-service training," "educational consultations" and "demonstration classes" that appear in proposals were devoted to business meetings, administrative assistance by teachers, and community work. Many agencies were ignorant of the services offerred by their "Regional Training Office."

---

[5] Presently all but two teachers have new directors.

[6] The participant without a degree is now the director of her agency's program.

Teachers felt that some bodies of knowledge concerning research with children and techniques used with handicapped children should be mandatory. Teachers wanted to know such things as "What are the language deficits of ghetto children, and if they exist, how specifically can I help my class?" "How as a teacher do I ascertain what will be expected of my children?" and "Who is responsible for interpreting my activities to the public schools and the community?" "Who is my support?" "If I as a teacher work with parents and community leaders, teach children, train my aide, and attend business meetings, who in the OEO framework looks out for the teacher?" This point (in many forms) was constantly reiterated throughout the seminar. "Who cares about me?" "Am I dispensible?" "Am I reflecting the frustrations of many parents?" "If Congress doesn't care, there won't be a program." A phrase "the poverty program syndrome" was an "in group" joke that served to express the group's anxiety.

Some teachers had previous experiences which they felt had helped them to "survive." These included group and agency work, enough experience with preschoolers to be confident in teaching activities, community acceptance, hearing a well-informed speaker discuss the problems of their agency community, and past employment or participation in the community in which they now taught. All felt they needed to know more--more knowledge concerning the unique problems of children, more understanding of sociological problems in the community and more concerning OEO as an orgainzation. All teachers expressed a desire to be observed, trained, supported and supervised. All wished to be recognized and given a voice in a program that they were asked to defend. "You have to believe in Headstart to teach in it."

Teacher Assumptions: How they effect curriculum? This area was examined from the view that what a teacher spends time on in the classroom reflects what she as a person feels children need to learn. When we described what teachers spent their day doing--one couldn't help feeling that perhaps their classes were being conducted similarly to any average preschool nursery. This observation disturbed the teachers! Upon examining the activities in each class, two activity areas were shared by all: 1) Routines: including toilet, milk, entering and leaving, rest and lunch. This consumed as much as 50% of each day. 2) Free Play: was the activity that teachers felt was necessary as a vehicle for many developmental activities in their classrooms. The time this activity consumed varied from 10-40% of the day. However, "free-play" in itself deserved a conference! Other activities were: arts and crafts, games, trips, story time and circle activities. The parent-teacher activities, teachers felt, often addressed themselves more realistically to child oriented problems than did the classroom.

Upon examining classroom programs and the assumptions behind them--the group experienced shock and frustration at the distance between the activity in the classroom and the needs of the children. Teachers could not help feeling that their programs reflected those models developed for "white, middle class" children--and that perhaps "we value these in such a way--we don't want our classes to miss this." The feeling of "making up" for something missed seems very often to be the essence of the Poverty Program. Although teachers could communicate and illustrate the physical and intellectual needs of children in their classes, classroom activities often did not reflect these needs of children. "If the problems of these children are so unique then why isn't the program?" Teachers did not have any way of evaluating the impact and outcome of their teaching--so it never changed. This kind of isolation seemed to be attributed to a lack of specific training with culturally deprived children and an absence of an evaluation program to describe public school expectations (and inequities if they exist).

The public school, their involvement in the Poverty Program, their hostility toward it, their ingnorance or indifference to it, all became areas for much discussion. Teachers felt that the public schools' inability to deal with the special problems of children and their "unwillingness" to develop a realistic curriculum was the real justification for programs like Headstart. However, this too became clouded with doubt when one teacher said "is it really fair to give children the feeling that Headstart is like school--maybe they'd be better off never given this freedom." All but one teacher felt the schools were inadequate. All reported hostile reactions toward the Poverty Program by many public school teachers. No Headstart teacher had ever been part of a meeting including Headstart personnel, kindergarten or grade one public school teachers. Headstart teachers saw public school personnel as unsympathetic to the "special" needs of Headstart children and that children were often penalized for their participation in Headstart." Parents, too, were reported to be suffering the contrast between "total acceptance in Headstart and the negative indifference of the public schools." Teachers in Headstart expressed a desire to observe in public schools--and to have kindergarten and grade one teachers observe their children prior to their entrance in public schools. A positive desire to cooperate in exploring ways to bridge the gap between Headstart and grade one was expressed.

What is Headstart All About? Headstart has been defined separately by children, parents, taxpayers, and teachers. It is seen differently by all. The seminar simplified this by attempting to isolate teacher goals for children. The following are a sample of aims expressed during the discussion. The purpose of Headstart was: 1) To develop a positive self image within children. 2) To give children an opportunity to socialize. 3) To give children an understanding of their community and it's vocabulary. 4) To give children an understanding of the vocabulary and surroundings of more advantaged children. 5) To prepare children for public school intellectually and behaviorally 6) To prepare parents for intervention with their children. 7) To improve the mental and physical health of preschool children. 8) To involve the community in the educational fate of their children. 9) To develop new ways for remediating learning deficits. All goals were not held in common.

The lack of uniform program impact was attributed to lack of communication within agencies, agencies with communities and agencies with each other and the OEO administration. Teachers saw a need for realistic training on the part of agencies working with disadvantaged children. Supervision and support by people specifically trained in the area was also cited as a need. Lastly there needed to be a separation of goals for teachers, parents, politians and agencies. Impact varied because of the wide variety of needs being administered to with the same "OEO dose."

The wide spectrum of problems facing Headstart teachers and administrators could well be discussed on a continuing basis, with an attempt to deal with the information in a scientific way. However, with this project, when the tapes were edited, "positive" and "negative" were clearly a matter of opinion. The participants themselves felt that the "tapes" spoke for themselves and that we should "tell it like it is." The problems are constantly changing. This fluctuation, the biggest aggravation, is perhaps the program's salvation. "It allows us (teachers) to do what we want before we are told what we have to do." The indecision which the group felt at the terminating session of the seminar exemplifies the indecision of the program's future.